

LIBRARY OF THE
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

510.84

Il6r

no.493-498

cop.2



170.87
Illbr
70 497
Cap 2

Math

UIUCDCS-R-72-497

COO-2118-0029

Parallel Image-Processing For
Automated Cervical Smear Analysis

By

John S. Read

October 1972



THE LIBRARY OF THE

NOV 9 1972

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN • URBANA, ILLINOIS

COO-2118-0029

UIUCDCS-R-72-497

Parallel Image-Processing For
Automated Cervical Smear Analysis

By

John S. Read

Department of Computer Science
University of Illinois
Urbana, Illinois

This work supported in part by Contract AT(11-1)-2118 with
the U.S. Atomic Energy Commission.



Digitized by the Internet Archive
in 2013

<http://archive.org/details/parallelimagepro497read>

ACKNOWLEDGMENT

I wish to acknowledge the assistance and support given to me by several people while I was working on this thesis. Professor B. H. McCormick, my adviser, provided endless patience as well as good advice. My wife and daughters gave up many months of family-type activities. S. N. Jayaramamurthy, Val Tareski, and Peter Raulefs wrote programs to implement Michalski's and McCormick's ideas used in the texture analysis section. S. N. Jayaramamurthy also collaborated in a paper which was incorporated in this thesis. Mrs. Judy Arter and Mrs. Patti Welch typed the several revisions of the text. I also wish to express gratitude to Dr. Ben T. Williams and Dr. Fernando Toledo, who supplied encouragement and indispensable information about cytology, and to give special thanks to Mr. Lawrence Brady, Cytotechnologist, who gave up his Thursday evenings for several months to teach me some of the rudiments of his art.

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION.	1
A. <u>Image Processing in Automated Cytology</u>	1
B. <u>The Subject and Scope of this Thesis</u>	2
II. CERVICAL SMEARS	7
A. <u>Purpose and Use of Cervical Cell Sample</u>	7
B. <u>Image Characteristics of the Cell Sample</u>	8
1. <u>Variations in Epithelial Cell Morphology</u> <u>Due to Differentiation</u>	8
2. <u>Model of the Cell Sample</u>	9
C. <u>Characterizations of Malignant Cells</u>	16
1. <u>Morphological Characterizations</u>	16
2. <u>Non-Morphological Characterizations</u>	20
III. REVIEW OF CYTOLOGY AUTOMATION	24
A. <u>Applications</u>	24
B. <u>Specimen Transport</u>	24
1. <u>Microscope Slides</u>	24
2. <u>Fluid Transport</u>	26
3. <u>Linear Deposition</u>	28
4. <u>Other Systems</u>	28
C. <u>Measurement and Analysis</u>	28
1. <u>Morphological Analysis</u>	28
2. <u>Non-Morphological Analysis</u>	42

	<u>Page</u>
IV. EQUIPMENT AND PROCESSES USED.	45
A. <u>Programmable Scanners</u>	45
B. <u>Show-and-Tell</u>	48
C. <u>Parallel Image Processing and PAX II</u>	48
D. <u>Texture Recognition using Varivalued Logic</u>	50
V. CYTOLOGICAL IMAGE PROCESSING EXPERIMENTS.	61
A. <u>Blob Detection</u>	61
B. <u>Cell Detection</u>	68
C. <u>Texture Processing</u>	71
VI. CONCLUSIONS	88
A. <u>Blob and Cell Detection</u>	88
B. <u>Texture Analysis</u>	89
C. <u>Recommendations for Further Work</u>	92
LIST OF REFERENCES.	94
APPENDIX.	114

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1.	Cross-section of Stratified Squamous Epithelium (left) and Views of Representative Cells.	10
2.	Photomicrograph of Epithelial Cells and White Blood Cells.	11
3.	Cell Structure as Delineated by Papanicolaou Stain . .	12
4.	Single Squamous Cell Status.	13
5.	Squamous Cell Interactions	14
6.	Photomicrograph of Cervical Smear Showing Cell Interactions.	15
7.	Block Diagram of Equipment Used.	46
8.	One-dimensional Texture-processing Example	
	(a) One-dimensional "Textures," Quantized to Four Gray Levels.	53
	(b) Receiver Operating Characteristic Corresponding to Table 5.	56
	(c) Generalized Logic Diagram with Interval Covering of T^1 against T^0	56
9.	Input Images	
	(a) Photographic Image: Blobs.	65
	(b) Digitized Image: Blobs	65
	(c) Photographic Image: Cells.	72
	(d) Digitized Image: Cells	72
10.	Operations	
	(a) "BLOB" Operation, Three Gray Levels. IRADUB = 2, IRADLB = 1, IDROP = 2	73
	(b) "RADAR" Operation	74

<u>Figure</u>		<u>Page</u>
11.	Blob Detector Processing of Figure 9(b)	
	(a) Large Blobs Detected.	76
	(b) Blob Detector Output.	76
12.	Cell Detector Processing of Figure 9(d)	
	(a) Smoothed Image (KPOW = 2)	77
	(b) Magnitude of Differences in Y-Direction	77
	(c) Blobs Detected.	78
	(d) Cytoplasmic Borders in the X-Direction Detected	78
	(e) RADAR "Beams"	79
	(f) Cells Detected.	79
13.	Texture Analysis	
	(a) Photomicrograph of Drying Artifact.	81
	(b) Photomicrograph of Cell Nucleus Showing Chromatin Texture (Center).	81
	(c) Test Set of Texture Samples	84
	(d) Test Set with Hit Counts	85
	(e) Test Set Classified	86
	(f) Receiver Operating Characteristic for Chromatin vs. Artifact.	87
	(g) Receiver Operating Characteristic with Tapered Quantization of Samples	87
	(h) Hit Count Distribution.	91

LIST OF TABLES

<u>Table</u>		<u>Page</u>
1.	Some Parameters Characterizing the Present Cervical Smear Screening System.	8
2.	Measurements on Abnormal Cells by Reagen and Wied. . .	19
3.	Automated Cytology Applications.	25
4.	Flow Systems--Non-Morphological Measurements	44
5.	Statistics Derived from the One-dimensional Texture of Figure 8(a) using a 1x3 Template to Define Events .	53
6.	Blob-detector Performance as a Function of Small-blob Parameter Values	67

I. INTRODUCTION

A. Image Processing in Automated Cytology

Images of biological cells are being increasingly used in quantitative studies of cell properties. New techniques have become available which elucidate subtle cell structures and functions. Procedures such as stoichiometric staining, fluorescent or radioactive tagging with antibodies or chemical precursors, and autoradiography produce images whose optical properties can be used to analyze complex biological events. However, conversion of optical properties to numeric form and subsequent interpretation become problems, especially when large cell populations are involved. Information in biological images is typically rather "noisy," owing to the heterogeneity and variation inherent in such materials. Simple automatic approaches to data extraction are sensitive to this noise. Frequently, morphological considerations are needed to limit the domain of measurement to particular structures to avoid spurious signals [97, 120, 191, 261]. Data conversion and interpretation by humans, on the other hand, can cope with the noise problem, but are not cost-effective for some applications of interest, even when augmented by interactive computation facilities.

What is needed, then, is the development of cost-effective image-processing algorithms which can cope with morphological criteria at the level of complexity present in biological materials. These algorithms can produce some measurements directly, or can be used in conjunction with other analytical instruments. For example, an image-processing algorithm could be used to locate a cell nucleus, whereupon an electron microprobe could be automatically directed to perform a chemical analysis. Conversely, image-processing could be used to automatically apply quantitative adaptations of human-oriented morphological criteria to

the output of existing high-speed cell analyzers. Flow and electrostatic particle transport systems have the capacity to handle, analyze and even sort large numbers of cells, often as many as 100,000 per minute. (See section III.B.2.) These devices typically apply measurements, such as light absorption at a particular wavelength, to an entire cell at once with no attempt to resolve intracellular structure. Some applications would benefit from the ability to rapidly filter large quantities of cells using high-speed whole-object measurements followed by automatic morphological analysis of the residue.

B. The Subject and Scope of this Thesis

In this thesis a particular application of digital image processing in cytology is examined: automated analysis of the well-known "Pap" smear. Pap smears are samples of epithelial (skin) cells from the uterine cervix and are used to detect cancer of the uterus while it is in an early stage and relatively easy to cure. Development of an automatic device to screen these samples has been recognized as of great potential value, almost from the time the test was first devised by Papanicolaou. However, this has proven to be a formidable task, as demonstrated by unsuccessful development projects undertaken by such organizations as the National Cancer Institute, IBM Corporation, and Vickers, Ltd. Accordingly, the goals of this thesis are prudently restricted to something less than a complete implementation. Actual implementation of a system to analyze cervical smears would require a careful analysis of the options available in staining, specimen preparation, transport and sensing to determine the most effective combination. It seems likely that much deeper insights into the nature of cancer cells will be forthcoming in the next few years. One would expect that this new knowledge will have a significant effect on the technological direction.

As is discussed in some detail in the next chapter, a cervical smear consists of a mixed population of cells spread on a microscope slide and stained. The cells have a tendency to clump together, fold, and otherwise present a very heterogenous and confusing picture to an observer, whether human or machine (see Figure 2). This clumping and overlapping will frequently be a problem whenever loose cells are to be examined, even if special preparative techniques are used to disperse them. The fact is that many cell types are supposed to stick together, and treatment of the sample to defeat this may result in the loss of important information.

In the present work, the prime objective was to examine the effectiveness of a parallel digital image processor in the analysis of cervical smear imagery. A wide range of activities could support this purpose. In order to define the problem adequately for machine implementation, and also to avoid expending excessive time on a complex biomedical problem, it was decided to accept as correct certain conclusions drawn in the course of development of the Cytoanalyzer project, an early attempt to automate the screening of cervical smears by image processing. The Cytoanalyzer project is discussed more thoroughly in chapter III. The pertinent conclusions, paraphrased from references [206] and [207] are as follows:

(1) Normal cells shed from the cervix and vagina exhibit a functional relationship between size and optical density of the nucleus that approximates $e(n) = Q/D(n)^2$ where $e(n)$ is nuclear optical density and $D(n)$ is nuclear diameter suggesting that the nucleus contains a constant quantity, Q , of dye-binding material in normal cells throughout various stages of differentiation.

(2) There is a continuous spectrum of change toward increased optical density of the nucleus and increased nuclear size (relative to

overall cell size) as one progresses from normal cells to cancer cells.

(3) Smears classified as being associated with cancer show the presence of a second population of aberrant cells superimposed on the normal cell population.

These conclusions were based on the use of a standard Papanicolaou-prepared smear, and were qualified by the statement:

"The application of these methods to an automatic instrument implies the condition that the instrument make no errors in measurement or cell recognition." [206, p. 468]

As is described in Chapter III, a prototype instrument incorporating these principles was built which was intended to measure nucleus optical density and nucleus diameter of epithelial cells. Due to limitations in the image processing technology then available, it proved to be impossible to meet the conditions of the qualifying statement. The prototype could not distinguish enlarged, dark, epithelial nuclei from clumps of white blood cells nor distinguish white blood cells from certain normal epithelial cell nuclei. Thus, the validity of the conclusions reached in the design study could not really be tested in a clinical environment because of the presence of this biological image "noise."

The experiments reported in this thesis were motivated by the assumption that improved image processing technology, as represented by a parallel digital image processing device (described briefly in Chapter IV) will permit reliable analysis of much more complex patterns of cell images. The emphasis here is not on the discovery and extraction of parameters for distinguishing a malignant cell from a non-malignant cell, but rather on the construction of algorithms for a parallel processor which permit the machine to rapidly make sense out

of the mess of cells and debris in the microscope field so that subsequent measurements (whatever they might be) are made on the correct objects: epithelial cell nuclei, for example, rather than clumps of white blood cells, cytoplasmic folds or other locally similar-appearing phenomena.

The interest in algorithms for a parallel image processor stems from the belief that for this and similar applications, parallel digital image processing has the greatest likelihood of being sufficiently fast, flexible and cost effective to perform the critical initial steps of object location and identification.

The remainder of this thesis consists of two parts. The first, Chapters II and III, contains background information about the application and a review of the rather extensive literature on automated cytology. The second part, Chapters IV-VI, describes some experiments in applying parallel image-processing to some of the cervical image difficulties mentioned above. The objective was to write and test programs for hardware like the Pattern Articulation Unit of Illiac III, where the programs could cope with three particular aspects of the analysis of cervical images:

(1) Rapid filtering of images of minimal photometric and spatial resolution to detect dark blob-like regions representing potential malignant cell nuclei. The Cytoanalyzer study mentioned above provides the justification for searching for large, symmetric dark regions. The constraint to use images of low resolution is in keeping with the realities of scanner performance in situations where high speed is required. The filtering procedure is designed to distinguish between dark blobs caused by symmetric clumps of leukocytes and blobs caused by other objects, including malignant cells.

(2) Detection and counting of normal, well-differentiated epithelial cells. This is again a filtering operation to be done on a low-resolution image, and is useful in cervical smear analysis as an indication of cell sample adequacy and as an input to a stopping rule.

(3) Textural discrimination between cell nuclei and blobs caused by drying artifacts. After the blob filter isolates areas of potential interest, blobs caused by spurious phenomena such as drying artifact can be eliminated by texture or spectral analysis on a higher resolution image. Data rate requirements are much reduced by rescanning at higher resolution only areas which require higher resolution to clear up ambiguities.

While there are many other aspects of cervical cell image analysis which could be of interest, it is felt that these three are particularly compelling, since previous attempts to handle them were unsuccessful because of the state of image processing technology. By capitalizing on the recently-available ability to make better use of two-dimensional information at high speeds, it is felt that these difficulties are solvable.

II. CERVICAL SMEARS

A. Purpose and Use of Cervical Cell Sample

The epithelial cells of the body cover and line organs, providing protection in some cases and secreting or absorbing vital fluids in others. Since epithelial cells are subject to various stresses, they are renewed and sloughed off (exfoliated) continuously. [63, p. 189-218]. In some cases, exfoliated cells are carried by fluids to locations in the body where they can be easily sampled (sputum, urine). These phenomena provide a painless and low-cost way of obtaining a pseudo-biopsy of tissues which may otherwise be accessible only by surgery.

The epithelium of the uterine cervix, being relatively accessible, can be sampled more directly by scraping. Under the present, manual system the cells are spread on a microscope slide, fixed, and stained according to procedures described by G. N. Papanicolaou in the early 1940's [156-158, 244]. An appropriately-trained person determines whether the sample contains tumor cells or cells from "pre-cancerous" conditions by looking for certain features of individual cells, and by looking at the overall pattern of cells on the slide. This task is commonly broken into two parts: a preliminary screening to eliminate indisputably normal slides, and then a close scrutiny of the remaining slides by a pathologist. The screening phase is usually delegated to a cytotechnologist, a technician who has undergone at least six months of formal training in cytotechnology followed by another half-year of on-the-job training. Some parameters which describe the effectiveness and load on this system are presented in Table 1, which is a compilation of information from several sources.

Table 1. Some Parameters Characterizing the Present
Cervical Smear Screening System

Number of deaths due to uterine cancer (US)	14 thousand/year [1]
Number of slides processed (US)	15 million/year [279]
Throughput capacity per screener	10 slides/hour
Time to screen one slide	2-5 minutes [165]
Cost per slide	\$1.00-\$3.00 [165]
Number of cytotechnologists (US)	4500 [68]
Percent incidence, various abnormalities	
Dysplasia	.5-.8 [173]
Carcinoma in situ	.3-.4 [173]
Invasive cancer	.1 [173]
Probability of missed positive	.1 [47,110]

While the primary purpose of taking Pap smears is to detect cervical cancer at a stage when it is easily controlled, other information is frequently obtained from the cell sample:

The quantity and types of cells present indicate whether the cell sample is adequate. For example, a sample consisting entirely of white blood cells was probably not made correctly.

Various micro-organisms such as yeasts, bacteria and trichomonads may be identified.

The degree of differentiation of the epithelial cells present gives an accurate indication of the status of hormonal activity. This is useful in estimating the effect of hormone therapy or detecting ovarian cancer in post-menopausal patients.

B. Image Characteristics of the Cell Sample

1. Variations in Epithelial Cell Morphology Due to Differentiation

Figure 1 shows on the left a cross-section through epithelial tissue like that present on the outer (vaginal) part of the uterine cervix. New cells are continuously being generated in the bottom (basal)

cell layer. In normal epithelium, as the cells differentiate, they become more and more flattened (squamous), and the nucleus shrinks. At the top (superficial) stratum, the cells are relatively easy to dislodge and can be harvested for examination by non-traumatic scraping or wiping. On the right in Figure 1 are top and side views of representative cells from various strata. Figure 2 is a photomicrograph of a Pap smear. Intermediate and parabasal cells are seen. The small dark blobs scattered throughout the field are white blood cells (leukocytes).

2. Model of the Cell Sample

As mentioned in the Introduction, the appearance of the cell sample in standard preparations is heterogeneous and confusing. To help sort this out, a model as depicted in Figures 3-5 was contrived. This model applies to mixtures of white blood cells (leukocytes) and to fairly well-differentiated squamous epithelial cells only, and shows some of the situations which occur. In this model, an idealized epithelial cell is postulated. Since the Papanicolaou staining process does not elucidate a living cell's true complexity, the cell can be regarded as having the parts shown in Figure 3. [63, 64].

Two classes of transformations act upon these cells to affect the image appearance. Various biological and mechanical forces can cause changes in the appearance of individual cells, as shown in Figure 4. In addition, interactions between cells produce the classes of cell images shown in Figure 5. Combinations of individual cell changes and cell interaction effects can produce complex images such as Figure 6.

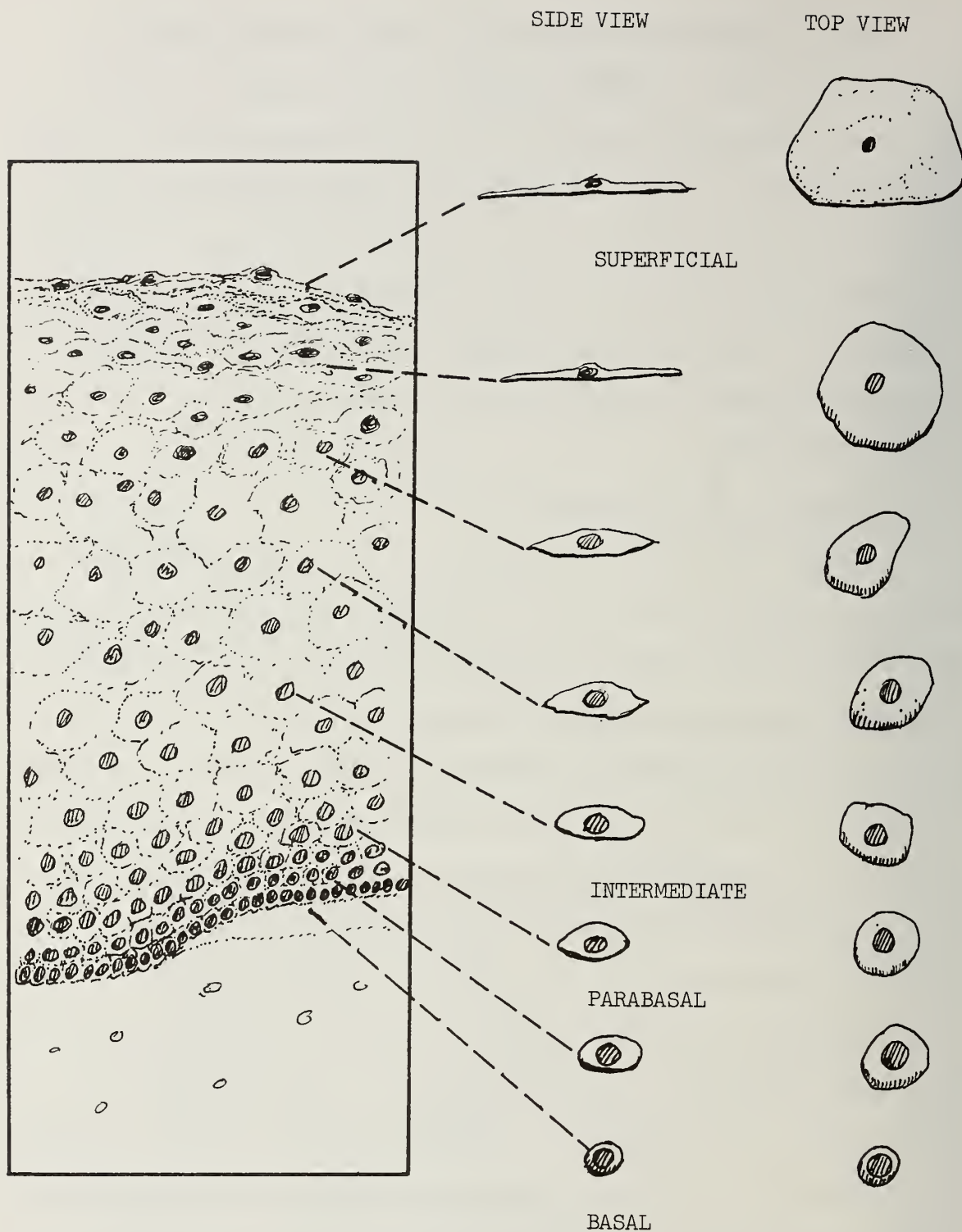


Figure 1. Cross-section of Stratified Squamous Epithelium (left)
and Views of Representative Cells

(Redrawn from [245])

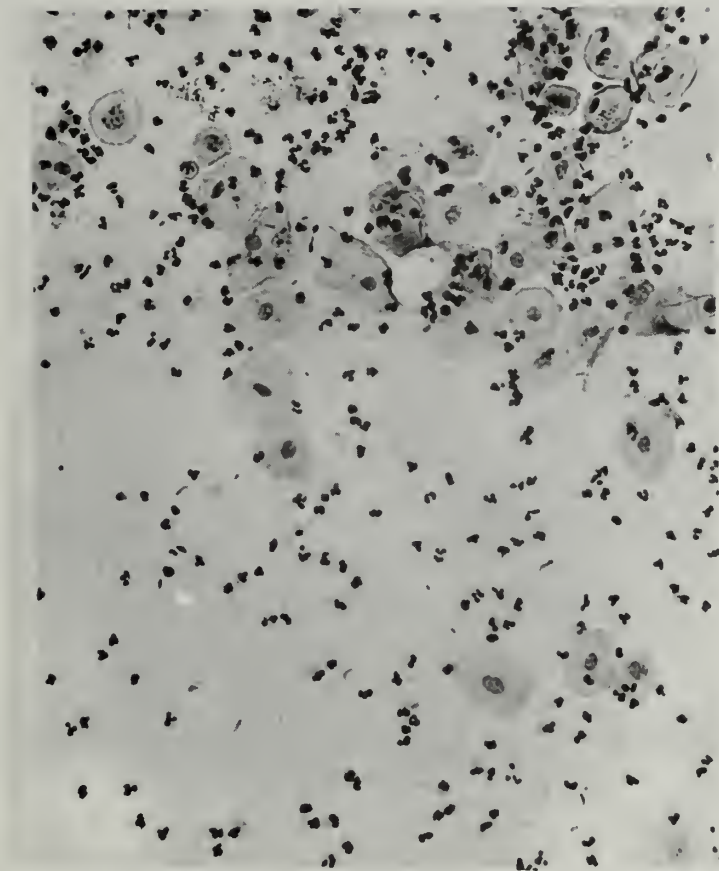


Figure 2. Photomicrograph of Epithelial Cells
and White Blood Cells

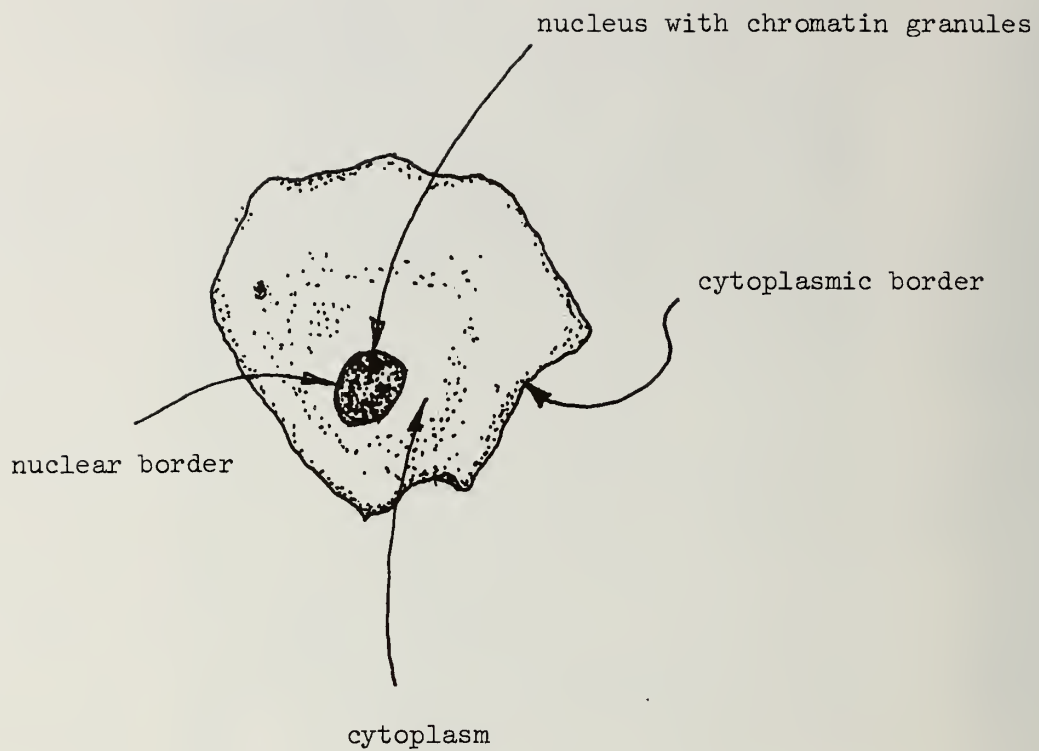


Figure 3. Cell Structure
as Delineated by Papanicolaou Stain


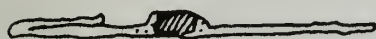
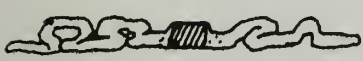
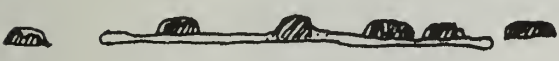

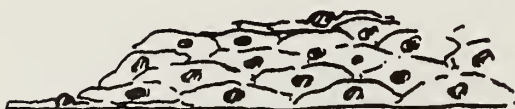
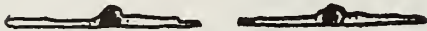
<u>CELL CROSS SECTION</u>	<u>CELL STATE</u>
	IDEAL
	FOLDED
	CRUMPLED
	WITH LEUKOCYTES
	CYTOLYSIS

Figure 4. Single Squamous Cell States

CELL CROSS SECTIONCELL STATE

IDEAL - NO CONTACT

OVERLAP OR TOUCHING
-NUCLEI NOT OVERLAID

OVERLAP - NUCLEI OVERLAID

STACKING

MASS

Figure 5. Squamous Cell Interactions

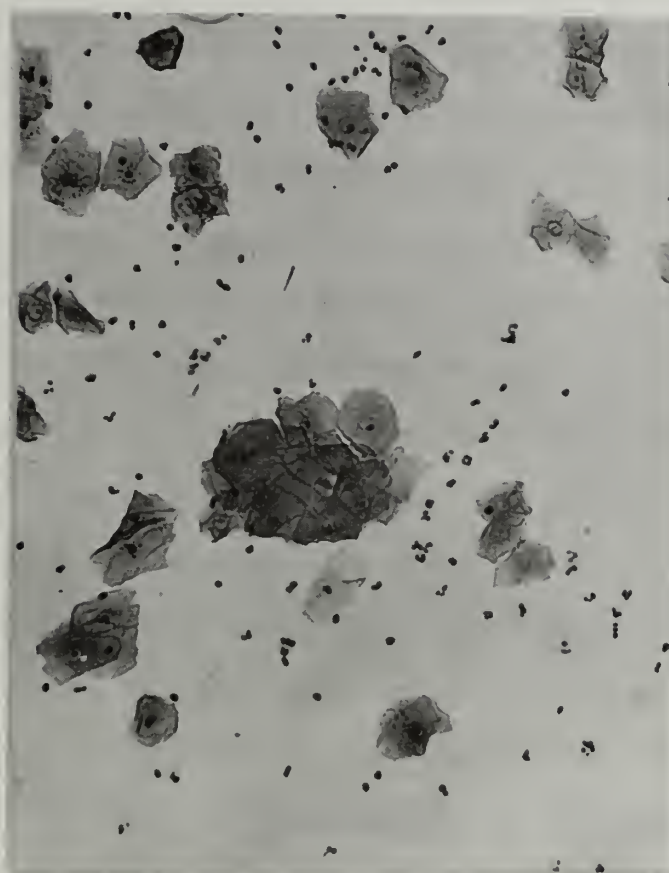


Figure 6. Photomicrograph of Cervical
Smear Showing Cell Interactions

C. Characterizations of Malignant Cells

A number of characterizations of malignant cells have been used or proposed for cancer cell detection systems; however, no characterization has yet been discovered which unequivocally applies always to cancer cells and never to normal cells [271]. As knowledge of cell structure and function deepens to the molecular level [64, 218], it becomes more likely that cancer cells can be characterized precisely in terms of their biochemical differences from normal cells. Once these differences are well-understood, development of instrumentation for accurate identification of cancer can be expected to follow rapidly. However, for the present, morphological parameters interpreted by the trained human visual system provide the only acceptable cervical cell analysis system [224] for cancer screening purposes.

The rest of this chapter contains a review of some of the criteria which have been proposed to distinguish cancer cells from normal cells. In some cases, machines have been built to make the necessary measurements, although none has yet been accepted for clinical use.

1. Morphological Characterizations

Under the present manual systems, cell samples are characterized in morphological terms; that is, in terms of form and structure as commonly interpreted by processes of human visual perception. The form and structure evident in Papanicolaou-prepared cervical smears is a consequence of the light-absorbing properties of the various stains used, and the differential uptake of those stains by the compounds comprising the cell substructures. In addition, there is a certain amount of modulation of the transmitted light by the cell material itself, as evidenced by its unstained appearance. Morphological identification of cancer cells is based on several criteria [52, 110, 156, 158, 173, 213, 245]:

- Cytoplasmic shape
- Nuclear area and shape
- Nucleus optical density
- Nuclear chromatin texture
- Ratio of nuclear to cytoplasm area

The cell-to-cell variability of these parameters is also considered to be highly informative in the sense that in cell samples from normal tissue, one cell looks very much like its neighbors. In cancer, this regularity and predictability is upset. For this reason, among others, it is frequently stated that the cytologist does not diagnose individual cells, but rather interprets the entire slide.

It must be pointed out that listing these criteria gives a false impression of the objectivity of the definition and use of these factors, as indicated by Spriggs [271] in a discussion of cytologic diagnosticians:

"If you want a simile, we are more like art connoisseurs, whose opinions are built up after years of studying the styles of different masters. The art expert cannot get much information from exact measurements, but he can identify the work of many different hands and he knows a lot about imitations. To the end of his life he is sometimes wrong, and nobody knows how often."

Some efforts have been made to quantify these impressionistic parameters. Wied et.al. [246] and Reagen, et.al. [173] measured large numbers of atypical cells in smears from patients with various forms of malignant and non-malignant conditions. Their results are presented in an abridged form in Table 2. Various conditions are listed across the top. Dysplasia is a non-malignant disruption of orderly growth and a loss of differentiation which occurs in either stratified squamous epithelium or in the mis-differentiated epithelium of metaplasia [173], and is characterized by cells with "sufficient" cytoplasm and much enlarged nuclei which have a fine-texture chromatin. Carcinoma in situ is an early form of cancer which eventually may go on to invade surrounding tissues. It is characterized, according to Table 2, by relatively

small cells (compared with normal superficial cells) having large nuclei with coarsely clumped chromatin. The various forms of invasive cancer have increased ratios of nuclear area to cell area, and coarsely granular or opaque, dark chromatin. In these two studies, no attempt was made to establish the discriminating power of the features. The objective seems rather to have been to confirm and objectify existing subjective criteria used in visual screening, without using the numbers to analyze the criteria.

Another study was conducted by Tolles, Horath and Bostrom [206,207] to establish a set of features to be measured by the Cytoanalyzer screening instrument. In this study an attempt was made to determine effectiveness of certain morphological features in detecting cancer. This research concerned three measurements: cell diameter, nucleus diameter and nucleus optical density. Two hundred cells were selected at random from each of over seven hundred slides. All isolated cells except white blood cells were measured for the three parameters by use of a manual microphotometer. The conclusions reached by Tolles, et.al. are described above in Section I.B. and are in basic agreement with Wied and Reagen.

Note that chromatin texture as a feature is ignored by Tolles et.al., even though it is considered to have great discriminating power by most authorities. [1, 52, 110, 173, 245]. This may have been because texture analysis at high speed was not then technically feasible. At any rate, as shown in [207], a screening performance of approximately zero false negatives (cancer smears called normal) vs. 15% false positives (normal smears called abnormal) could be expected for premenopausal women sampled by cervical scrape, using nucleus optical density and nucleus diameter as features. To obtain this performance, a smear was to be called abnormal if more than 1.5% of cells had optical density

Table 2. Measurements on Abnormal Cells by

Reagen [173] (R), Wied [246] (W).

	Dysplasia		Ca. in situ		R	Keratinizing	R	Non-keratin	R	Small	Normal	
	R	W	R	W							R	W
Cell area (μ^2)	1088 + 311 - 995	995 - 392	353 + 116 - 116	381 - 112	275 + 107		256 + 69		169 + 37		1600 + 1550	
Nuc. area (μ^2)	167 + 38 - 35	174 - 35	109 + 33 - 33	116 + 26	77 + 28		88 + 30		65 + 13		37 + 11	39
Rel. Nuc area (%)	16 + 43	17	32 + 6	31	29		34		39			
Isolated (%)	81		72	69								
Aggregates (%)	19		28	31								
Sheets (%)	19		3									
Syncytia (%)	0		25									
Finely granular (%)	94		22		22		20		13			
Coarsely granular (%)	4		72		59		76		80			
Translucent opaque (%)	2		7		19		3		8			

Arrangement of
Abnormal CellsChromatin
Texture

greater than .5 and nucleus diameter greater than 10 microns. Performance with these features was shown to be very sensitive to sampling technique, with scraping producing the aforementioned excellent performance and vaginal aspiration resulting in over 65% false positives for 5% false negatives. Vaginal aspiration is known to be very poor at harvesting cells from squamous cell cancer, especially carcinoma in situ [1, p. 57]. Actual performance realized by the Cytoanalyzer is discussed in the next chapter.

Quantification of chromatin texture analysis has been somewhat more difficult, since there are no universally understood measurements of texture. The very subjective texture characterizations of Table 2 illustrate this. Some progress has been made toward this goal, however, via experiments with computer-oriented texture measurements [15, 125, 166, 172, 187]. Some of these algorithms are discussed in Chapter III.

The parameters of cell-to-cell variability and cell arrangement have apparently not been subjected to objective analysis beyond that of Table 2.

2. Non-Morphological Characterizations

This section is included to provide context for the discussions of morphology and to provide background for the section on non-morphological measurements (III.C.2.). No claim is made for completeness or currency.

Much is being learned about the processes of malignancy which very probably will be applicable to automated cancer detection. Some of these insights may lead, for example, to development of more specific staining procedures which could vastly simplify an image-processing approach to cell analysis. Other criteria might lend themselves to non-morphological techniques such as automated chemical analysis or high-speed spectrophotometry.

Increased DNA Content

It has been shown [185] that cervical cell samples from women with cervical cancer typically have a population of cells with increased DNA content superimposed over the population of normal cells. An amount of DNA in a cell nucleus which corresponds to the amount found in a normal interphase (between divisions) body cell is called the diploid amount. The amount found in a replicating cell is, on the average, twice the diploid amount and is called the quadruploid amount. If a histogram is made showing number of cells vs. DNA amount, then a mode normally occurs at the diploid amount for cervical cell samples. If for some reason replicating cells are included, a mode will occur at the quadruploid level. In samples of cancer cells, the histogram is spread out, with large numbers of cells containing aneuploid amounts of DNA, and with modes occurring at places other than at diploid and quadruploid levels, indicating the presence of abnormal stem lines or clones [27, 72, 247]. DNA has properties which facilitate determination of its quantity photometrically. It absorbs ultraviolet light in proportion to its mass, especially if the wavelength of the light is in the vicinity of 260 nanometers. [267]. Also, DNA binds basic dye molecules. The mass of bound dye can be measured photometrically, permitting the DNA mass to be inferred [237, p. 2].

Immunological Indicators

It is a theory that the transformation of a normal cell into a cancer cell includes changing the chemical structure of the external cell membrane to the point that the cancer cell has antigenic properties like those of a foreign protein entering the body tissues from outside. In normal circumstances, the antibody

response would dispose of the tumor cell. In cancer, however, the immune system is somehow suppressed, and the tumor goes unchallenged. [218, p. 595]. Presence of tumor-specific antigen has been demonstrated in hamster cells transformed by a virus into cancer cells. [218, p. 611]. The antigen can be detected by the fluorescent antibody technique in which antibodies are tagged with fluorescent dye molecules. [254]. When the antibodies attach to the antigenic tumor cells, the fluorescent dye permits identification by fluorescence microscopy. [92]. Conversely, tumor cells in humans have been detected by measuring the absence of isoantigens. Isoantigens are antigens which cause antibody responses when cells are transferred among individuals of the same species, as in the case of the A, B and O blood types. Davidsohn, et. al. [38] claim to be able to demonstrate a progressive loss of isoantigen as a precursor to the spread of a tumor to noncontiguous tissue. The degree of loss of isoantigen is held to parallel the degree of loss of normal cell differentiation. Some related characteristics of cancer cells are also due to the alteration of the normal chemical composition of the cell's external membrane. For example, it has been hypothesized [218, p. 591] that animal cells stop multiplying when they touch one another (contact inhibition). Some cancer cells do not have this property, and continue to reproduce even when piled up, as demonstrated in cell cultures. Also, normal animal cells show "selective stickiness" i.e., they prefer to adhere to cells of their own type. Again, cancer cells do not show this, providing more evidence for the existence of changes in the cell membrane. This change can also be reflected in an altered electrical charge on the cell surface which can be measured by electrophoresis

[249]. It was also discovered that the surface changes result in a tendency for cancer cells to be precipitated by a certain glycoprotein derived from wheat germ [249; 218, p. 596]. Unfortunately, wheat germ agglutinin also appears to precipitate normal parabasal and some benign though atypical cells. [249].

Presence of Abnormal Cell Products

Another theory supposes that cancer can cause changes in the chromosomes of a cell. [218, p. 597]. If so, it would be reasonable to expect to find changes in the proteins coded by the chromosomes and in the biochemical processes controlled by those proteins. These phenomena have been observed and proposed as cancer-cell characterizers. One such discovery was that many types of cancer cells excrete more lactic acid than do similar normal cells. This was determined to be due to an unexplained change in the (unknown) mechanism which regulates glucose consumption. [218, p. 594]. It is also possible that production of abnormal enzymes can be detected by biochemical assay using an automatic analyzer [248] or by development of specific staining procedures [250].

III. REVIEW OF CYTOLOGY AUTOMATION

A. Applications

Research and development in automating the measurement of microscopic objects has at least a quarter-century's history. Early impetus to this work was given by a desire to automate the extraction of size distributions of a variety of micron-sized particles ranging from red blood cells to coal dust, since statistically significant estimates of these were and are very costly and difficult to obtain manually [217]. Recently, much work has been engendered by a desire to develop instrumentation for clinical use in the cytology laboratory, and for making automated, quantitative photometric studies of cell populations. A wide range of purposes can be discerned, as suggested by Table 3, a somewhat arbitrary sampling of work done in the field of cytology automation.

In the next sections, several cytology automation projects are discussed in terms of the techniques used to solve two problems which seem to be fundamental: presentation of the cells to the machine for analysis, and implementation of measuring and analysis strategies.

B. Specimen Transport

The method by which cell samples are prepared and presented to the sensor system is critical in determining feasible approaches to the analysis problem. Three fundamental approaches have been tried in the past: (1) microscope slides, (2) fluid transport, and (3) linear deposition.

1. Microscope Slides

Microscope slides are the traditional method of carrying cells for presentation to human microscopists and to a variety of instruments. Initial attempts to develop high-speed scanning devices used slide transport [129, 127, 115, 87, 199-203, 208]. However, this approach

Table 3. Automated Cytology Applications

Some Representative Bibliographic References

Detection of Individual Red Blood
Cell Abnormalities

Green [60]

Cell Sorting

Fulwyler [54]
Hulett [73]
Kamentsky [87]Red Blood Cell Count and Size
Distribution Abnormalities

Brecher [28]

Microbe Colony Monitoring

Glaser [57]
Mansberg [112]

Neuron Counting

Dudley [19]
Lipkin [107]
Mansberg [115]

Chromosome Karyotyping

Butler [277]
Castleman [276]
Ledley [102]
Mendelsohn [132]
Neurath [278]
Rutovitz [260]
Wald [216]

Detection of White Cell Abnormalities

Ingram and Preston [14]

White Cell Differential Count

Kamentsky [155]
Prewitt and Mendelsohn [166]
Preston [163]
Technicon Corporation
Young [242]

Exfoliative Cytology

Husain/IMANCO [261]
Ishiyama [76]
Kamentsky [94]
Nuclear Research Associates
[62]
Tetronics Corporation
[66]
Tolles [191]
Vickers Corporation [135]
Wied [227]

Live/Dead Cell Count

Kamentsky [87]

DNA Distribution

Van Dilla [211]

turned out to be infeasible because the sensing and analysis methods of the time could not distinguish between objects of interest and artifacts caused by clumping and overlapping of the cell sample while maintaining sufficient speed. Slide transport has been used successfully in instruments where manual location and centering of individual cells is permissible as in microspectrophotometric cell constituent analysis [7, 8, 27, 31, 185-187, 226, 228-232, 235, 237, 252], or in research environments where morphologic parameters for cell recognition are under investigation independent of speed requirements [13-18, 20, 60, 107, 125, 147, 161, 166, 168, 179, 227, 233, 242], or in those cases where the cell sample tends not to present severe clump or overlap problems [75, 102, 112, 163, 216]. However, there is motivation to use microscope slide transport in automatic cell analysis systems because (1) it greatly simplified interfacing machine results with traditional morphologic taxonomies, especially in the case of screening applications where positive samples need to be re-examined by a human. If the sample is disrupted by the transport technique, much valuable information about cell-to-cell relationships is lost. If a second sample must be taken in each case to permit human interpretation, intricate legal questions arise because of the possibility of missed positives in the second specimen. (2) In addition, a capability for interpreting slide-carried cells would permit the use of large numbers of existing labelled samples now stored in archives. For these reasons, attempts are being made to improve image-processing technology so that more complex cell images can be handled [123, 164, 255, 256].

2. Fluid Transport

Many of the problems in cell location, isolation and measurement are greatly simplified if the cells can be placed in a fluid (usually liquid) suspension and transported past the

sensors in an orderly, predictable manner. To implement this, the cell suspension has been transported via capillary tubes or flow cells, sometimes with special laminar flow conditions. The latter was initially described by Crosland-Taylor [37], and is usually referred to as a Sheathflow or Crosland-Taylor system [144, 211]. A capillary containing the sample suspension is enclosed in a larger tube in which a flow of clean fluid is maintained. As the sample suspension emerges from its capillary, it is surrounded by a sheath of clean fluid. The composite is drawn through a constricting nozzle which focuses the central stream down to a fine filament, perhaps only 10 μm in diameter. Clogging problems are reduced, since the constriction is done by the fluid sheath. Where clogging is not an acute problem, or if a close positional tolerance is not required, then tubes or channels of diameter on the order of 100 μm can be used.

Systems have been developed in which the cell population is physically separated into fractions on the basis of the measurements made. To do this, the solute stream can be broken into droplets by periodic vibration of the fluid column. A charge is applied to the droplets containing cells of interest, enabling them to be electrostatically deflected into a separating container. This technique was adapted by Fulwyler [54] from an ink-writing oscillograph of Sweet [262]. Another method of cell sorting is used by Kamentsky in his Rapid Cell Spectrophotometer [101]. A fluid switch causes selected cells to be pulsed into a side channel and saved on a filter, an approach motivated by a goal to select the one-hundred most diagnostic cells in a cervical cell sample for viewing under a single microscope field as a means of solving some of the re-screening problems in the use of flow transport systems.

3. Linear Deposition

This approach is similar to the fluid transport systems in that the cell sample is broken up and suspended in a fluid. However, instead of carrying the fluid past the transducers, the fluid is laid down in a narrow track on a plastic film or glass substrate using a device similar to a drafting pen. The track is subsequently carried past the transducers. This procedure has some of the advantages of both the microscope slide and fluid transport systems. The cells are flattened and static in shape, so conventional morphologic interpretation can be made relatively easily by machine or human. On the other hand, the object isolation problem is somewhat reduced. Track-laying systems have been developed by Vickers [135] and Tetronics [39, 66] Corporations in Great Britain for cervical cell screening machines. Both of these utilize a plastic film of about one meter in length for a cell sample, and lay a line of cellular "ink" approximately 1000 (Vickers) or 500 (Tetronics) μm wide. A scanner built by Mansberg for fluorescence measurements [116] included a long flexible glass ribbon to hold the specimen.

4. Other Systems

Mansberg also described a record-player like scanner in which a fluorescing sample on a membrane was scanned in a spiral pattern. Other specially designed transport systems for Petri dishes, etc., have also been considered [112].

C. Measurement and Analysis

1. Morphological Analysis

Morphology is defined by Webster as "the features comprised in the form and structure of an organism or any of its parts." Since form and structure are most commonly sensed visually, mechanized morphological measurements are typically made on some sort of simulation of a retinal representation of an image of the object. Morphology can be measured in other ways, such as by

measuring diffraction effects, but these involve fundamentally different techniques, and are discussed briefly in Section III.C.2.

To form the retinal representation, light energy transmitted by or reflected from the object is measured either by a retina-like array of sensors, or more commonly, a single sensor time-shared by scanning and sampling in a systematic pattern. This process generates a picture, defined by Rosenfeld [181] as a non-negative function of two variables which is non-zero only in a bounded region of some standard size and shape, e.g. a square. Two fundamental approaches have been used in making and manipulating picture representations, which can be named the one-dimensional and two-dimensional approaches. In a one-dimensional approach, the picture function is processed by sampling in a TV-like raster scan. This produces a time-varying video signal which can be analyzed by digital or analogue signal processing techniques transplanted from well-known engineering practice. A limited amount of two-dimensional information can be processed by the use of a shift register or delay-line buffer memory of, say, the previous few scan lines so that measurements of the same object on succeeding scan lines can be associated. This approach has been used in many high-speed microscope image analysis systems so far because fairly high processing rates can be achieved at relatively low cost. Processing occurs in real-time with the scan, which can occur at television rates. However, applications have been limited to those where a rudimentary analysis will work, e.g. where pre-set thresholds can be used to determine when an object boundary has been crossed.

Two-dimensional approaches do not constrain analysis to a particular spatial sequence. However, storage and accessibility of the image

become problems, as does processing time, since conventional number or character computers have a one-dimensional topology poorly adapted to picture analysis. Alternative architectures have been proposed [82, 106, 123, 257, 258], and a discussion of one of these is contained in the next chapter.

The rest of this section on morphological analysis considers various morphology-based systems which have been described for use in automatic cytology. Discussion will frequently be in terms of three processes: Object identification and isolation which involves locating cells and distinguishing them from other objects which might also be present, and also making object sub-segmentations; Shape analysis which involves parameters of object contours, however established; and Texture analysis, which is based on brightness measurements independent (except on a very local level) of shape information. There is some interaction of these processes, e.g., texture analysis can be used to establish a contour for making shape measurements, or can be used in object location. This breakdown of morphological analysis into shape and texture analysis is related to a definition of image by Huang, et.al. [259]: "We consider an image as the sum of three components: the low-frequency part, the edges and the textures." Texture is defined, therefore, as what is left over when the edges and low-frequency parts are subtracted out. Shape information is contained in the edge and low-frequency components.

In the Cytoanalyzer project, an attempt was made to develop a clinical instrument for cervical cancer pre-screening using scanning techniques transplanted and extended from particle counting and sizing technology [24, 25, 188, 199-205]. A one-dimensional approach was used, with hard-wired logic analyzing a signal generated by a Nipkow-disk image-plane scanner. Early design objectives [218] were to use Papanicolaou-prepared samples in order that the cytologist's re-examination

would be easier. A preliminary version included object-location circuitry which would find cells not overlapped by other cells by examining the sequence of signal levels as the scan crossed a cytoplasmic border, traversed the cytoplasm, then the denser nucleus, followed by more cytoplasm and finally dropping off the other border. The final implementation, [24], however, used a highly modified preparation, [36], which eliminated cytoplasmic staining and was intended to disperse the cells evenly on the slide. The image was scanned with a two- μ m aperture in an extended raster of 100 μ m x 5 cm. Each time the aperture traversed a chord of a nucleus, a video pulse resulted. The duration was proportional to the width of the chord and the maximum amplitude was proportional to optical density. These chords were summed, resulting in a measurement of nuclear area and integrated density, the parameters determined to be discriminating for cancer cells by the study discussed in II.C.1. Cell nucleus identification was accomplished by rejecting chord pulses preceded or followed by absorbing material. Since the cytoplasm was unstained, a true cell nucleus was supposed not to have this characteristic. Some rudimentary shape measurements were used to eliminate non-nuclei. However, a residual noise count of 10-50 abnormal counts was still experienced on a normal smear. Each nucleus was classified and counted as normal, abnormal, or indifferent. When the normal count reached 10,000, the machine stopped and the ratio of abnormal to normal counts (A/N ratio) was computed. If this exceeded a threshold, the sample was declared positive.

In clinical trials using aspiration-collected samples and a prototype Cytoanalyzer, a false negative rate of 10.3 per cent was achieved for a A/N ratio threshold which correctly identified 36.6 per cent of negatives. The false negative rate happens to be close to that estimated

as common and acceptable for human screeners (see Table 1), and the overall performance was not very much worse than that predicted for the vaginal aspiration samples [207]. However, these results caused the Cytoanalyzer prototype to be deemed unusable for pre-screening for two reasons [191]: (1) It was felt that the 10.3 per cent false negative rate was too high for prescreening, since the subsequent false negative rate of the human screener would increase the error to an intolerable level. It was felt that lowering the A/N ratio threshold to improve the false negative rate would force the false positive rate so high that the machine would not reduce work loads to a significant extent.

(2) Some of the counts registered were out of agreement with experimental data. An investigation led to the conclusion that the presence of large numbers of leukocytes overwhelmed both the normal and abnormal count categories because single leukocytes could not be distinguished from superficial epithelial cell nuclei and clumps of leukocytes could not be distinguished from enlarged, abnormal nuclei. Furthermore, individual abnormal cells were called indifferent or normal 82% of the time because of the placement of the decision boundaries.

An approach very similar to the Cytoanalyzer was taken in the "Automatic Cytoscreener" described by Ishiyama [76], although the objective was to develop a screening machine rather than a pre-screener. This device used glass slide transport, a CRT flying-spot light source, and one-dimensional image analysis with wired-in logic. The abnormal/normal decision for a slide was made on the basis of the frequency distribution of nuclear diameters. Nucleus identification capability seems to be absent, and no performance data at all on cervical material is given.

In a pioneering study, Prewitt and Mendelsohn [166] analyzed images of leukocytes using algorithms which have since seen use by several others [60, 163, 242]. An off-line CRT flying spot scanner, CYDAC, was used to generate two-dimensional representations of cell images which were later analyzed with a general-purpose digital computer. However, neither the object isolation nor the texture analysis described used spatial information directly. All processing used only the information contained in the optical density frequency distribution (gray value histogram) of the image. This seems to have been possible at least partly because the images contained only the cells of interest and were very low in noise content in terms both of random scanner noise and interference from other similar objects. Object location was confined to separating the smoothed histogram into three regions: background, cytoplasm and nucleus. This separation was accomplished by assuming that density value changes rapidly as the border between, say, cytoplasm and background is crossed. Therefore, these border density values will occur comparatively infrequently in the picture, and will show up in the frequency histogram as minima. The local minimum of lowest optical density value is taken to represent the threshold distinguishing background and cytoplasm picture elements and the minimum of highest density defines the nuclear boundary. These thresholds can be used to segment either the digitized picture itself or the density histogram derived from the picture. Rudimentary shape and texture features such as cytoplasm area, average nucleus density, density distribution skewness and several others are extracted from the density histogram, and classification occurs on the basis of techniques of standard statistical decision theory. The initial study reported in [166] only classified four of the five leukocyte types (basophil omitted)

and a training set of only 22 cells was used. There was no report of trials using an unknown set. Further experiments were reported in [167].

In contrast to the methodological and research orientation of the Prewitt-Mendelsohn project, the Perkin-Elmer Corporation's Cellscan/GLOPR system [74-75, 161-164] was intended to lead eventually to "a practical system for general use" [163]. The initial purpose [161] was to demonstrate the feasibility of semi-automatically scanning blood smears to locate very rarely-occurring binucleate lymphocytes, which are evidence of low-level radiation damage. The approach used is also in contrast to the Prewitt-Mendelsohn procedure in that optical density information is eliminated as quickly as possible and processing is done on a binary picture. On the other hand, later versions of Cellscan utilize a histogram technique to derive the binary pictures, segmenting the multiple density input image into cytoplasm/red cell and leukocyte nucleus picture elements. These planes are processed by a special purpose hardware image processor, the Golay Logic Processor (GLOPR) [258, 164]. The two-dimensional GLOPR operations can extract features for shape and texture analysis using local homogeneous operations very similar in spirit if not in implementation to PAU operations of the Illiac III (see next chapter). Texture analysis of the nucleus fine structure is accomplished by an iterative thinning algorithm which strips ones from the borders of connected components of the binary pictures. At each iteration, the number of remaining isolated ones are counted. A record of these counts as a function of iteration number is a texture feature which indicates the number and size distribution of granules in the nucleus [75]. Several shape measurements can also be made by using similar combinations of marking, propagation, and counting processes, for example to measure area vs. perimeter, or to locate and size concavities and inclusions. A late version of Cellscan/GLOPR

includes an automatic vibrating-mirror scanner with a hardware object locating capability. Apparently, the latter is done by one-dimensional wired-in signal analysis using simple criteria for identifying objects likely to be white cells. All versions of Cellscan have used microscope slide transport and standard staining procedures. Preparation has included spinning the slide to produce a monolayer of well-dispersed cells. The system can apparently do a differential leukocyte count as accurately as a human scanner [264]. However, two factors mitigate against its acceptance for clinical use: (1) As currently implemented, it is much too slow (overnight to do one slide); and (2) Technicon Corporation, the IBM of clinical lab equipment, recently introduced Hemalog-D, a flow system with non-morphological analysis, which can do the differential count at realistic rates.

The use of spectral (color) information has been investigated by Young [242] at MIT as a means of object isolation in color photomicrographs of blood smears. The objective is to classify each picture element as background, red cell or white cell. A color transparency is scanned with a broad-spectrum CRT-generated flying-spot. Two dichroic mirrors separate the transmitted light into red, blue, and green components which are each sensed by a photomultiplier and digitized to eight bits. The color data is encoded as two numbers, $r = R/(R + G + B)$ and $g = G/(R + G + B)$ where R, G, and B are the three digitized PMT outputs. The vector (r,g), called a chromaticity pair, is used as a feature vector for a maximum-likelihood classifier in making the red-cell/white-cell discrimination. A Prewitt-type histogram approach is also used to separate background points from cell points.

Husain [261] conducted a study to determine the best color filter to use in a Quantimet monochromatic image processor (see below) to

establish a brightness threshold capable of reliably separating nucleus from cytoplasm and malignant from non-malignant nucleus. His conclusion was that "density alone cannot work satisfactorily" because of an "unacceptable degree of overlap in some cases." Husain conducted a further study which generally confirmed the earlier Cytoanalyzer parameter analysis of Tolles, et.al.

Green used the histogram method to isolate and segment scanned monochromatic photomicrograph images of red blood cells in an attempt to develop quantitative morphologic measures equivalent to hematologists' evaluations [60]. Green used noisy images of much reduced resolution as compared to Prewitt-Mendelsohn's, and found that noise and quantization errors required a more complicated method of locating peaks and valleys in the histogram. A global histogram segmentation was followed by local histogramming and segmenting of the red cells. Object recognition was by fairly simple area and perimeter vs. area criteria. A large number of shape and texture features were extracted, including area, total optical density, and eccentricity. The primary objective was to extract these quantitative shape measurements, rather than to arrive at a diagnostic decision.

The Vickers trace-laying system mentioned in Section III.B.3. was intended to be used with a very simple optical device for locating dark blobs of size greater than 12 microns in cervical cell samples [135]. A fair amount of work was done to show that this would be an adequate parameter for pre-screening [175]. However, this claim seems to have met with massive skepticism [224, 169]. A more sophisticated texture and shape analysis was proposed by McMaster [125], with the 12-micron criterion to be used for quickly locating objects of interest [251]. Low noise, high resolution pictures digitized from photomicrographs of

individual, isolated cells were used. Five parameters were extracted: average transmission, average transmission difference in adjacent picture elements, average difference in successive centroid-nuclear border radii, average radius, and ratio of maximum to minimum radius. A product, P , of weighted functions of these parameters was formed, with the weights computed by maximizing the difference of the extreme values of P for a small training set of 18 normal basal cells and 11 malignant cells. No effort was made to try the classifier on any cells not in the training set.

Similar measurements were automatically extracted from photomicrographs of cervical smears by a system of Rosenberg and Ledeen [179]. Object location and identification was accomplished by a raster search stopped by a preset threshold density, followed by size and shape analysis. Measurements were extracted (average nuclear radius, average deviation of radii, nuclear area, average nuclear density) and conditional probability distributions computed for 100 normal and 100 malignant cells (pre-identified). Means and standard deviations agreed closely with Reagen and Wied (see Section II.C.1.). However, conditional distributions published in [179] show considerable overlap between normal and malignant cells, suggesting that the features extracted are not very good for cancer detection. This is rather interesting, considering that these measurements supposedly reflect criteria used by human screeners. Also, Rosenberg and Ledeen fail to mention whether the cell samples were from more than one person. No classification of unknowns was attempted using these conditional distributions.

A series of instruments for sizing and counting of various microscopic objects was described by Mansberg [111-116]. All used a one-dimensional approach, and assume that objects are fairly predictable in

density, well separated and not too complex in shape. A prototype scanner was described which was intended to scan an entire sectioned human brain (6000 whole-section slides) in 1200 hours to attempt to establish a quantitative correspondence between brain lesioning and neuron depopulation. To do this, it is necessary to distinguish glial cells from neurons. Since shape analysis is difficult with one-dimensional processing, this object identification problem was attacked by trying to defocus the spot enough to eliminate detection of the smaller glial cells. However, problems were encountered in controlling the spot size variation, and at last notice (1964), the glial cell/neuron discrimination problem had not been solved [113]. Mansberg also described a system for counting fluorescing biological objects stained with the fluorescent antibody technique [116] (see the previous chapter).

Another scanner for making fluorescence measurements on slide-transported smears was proposed in 1951 by Mellors, Papanicolaou, et.al. [127-130] for automating cervical cancer detection.

The Quantimet, a vidicon scanning instrument with one-dimensional analysis was developed by Image Analysing Computers, Ltd. for counting and sizing microscopic inclusions in metallurgical specimens. The same device was used for counting goblet cells in specially-stained sections of lungs of rats exposed to sulfur dioxide. However, the machine counts were not directly comparable to manual counts, since the Quantimet could

not distinguish stained goblet cells from random blobs of stain [120]. A more sophisticated version of Quantimet was used in some experiments to determine its usefulness in cervical cancer screening [255, 261]. Slide transport and standard Papanicolaou staining was used, although there was a desire to couple the Quantimet to the Tetronics trace-laying machine [261]. Newer versions of the Quantimet include an ability to extract several morphological features such as integrated density, perimeter, or presence of second phases (nuclei).

Both morphological and non-morphological analysis occurs in the Cytoscreener of Nuclear Research Associates [62] in which a programmed Ultraviolet CRT flying-spot scanner generates a raster on cells transported by a laminar flow system. The system measures total UV absorption as a measure of DNA content, and also analyzes nuclear size, nuclear density, cell symmetry and nuclear-cytoplasmic ratio using "a pattern-recognition computing system" not further described. The Cytoscreener also has an object-location capability in that it can recognize and ignore fragmented or clumped cells or debris. All this occurs at a rate of approximately 5000 cells per minute, which is rather slow, since a cell sample may contain 100,000 cells or more. Performance of the Cytoscreener in a clinical trial using cervical cell samples, was fairly impressive. One hundred specimens were used, of which 24 were known to be from cancer patients, and 76 were known to be from normal patients. The specimens were analyzed by conventional cytology and also by the Cytoscreener. The false negative rate was 8.3% for the Cytoscreener and 12.5% for the humans. However, only 22 of 76 normals were screened out by the machine, partly because 31 of the 76 were rejected as having insufficient cellular material for processing. (Compare with Rapid Cell Spectrophotometer performance, next section.)

George Wied and his associates have conducted extensive investigation into image parameters to distinguish malignant and normal cells. According to Wied [363], "there are two roots to the application of pattern recognition principles to biological cells. One has its origin in development of methods and instruments for quantitative cytochemistry... The other root... is found in commercial and academic interests for automating clinical microscope screening procedures."

Wied has developed a cell recognition system (TICAS) very strongly based in quantitative cytochemistry. Cells are located manually and scanned with a very slow scanning microphotometer (17-20 msec per sample point minimum). Three objectives are enunciated [238]: (1) Description and discrimination of cells which are known to be biologically different but which are difficult to distinguish by standard techniques, (2) Providing computerized access to images of cells accompanied by expert diagnostic opinions, (3) Computer-aided instruction in cyto- and histopathology. A system of remote access to TICAS via phone lines has been proposed [238] and now implemented, wherein an image generated by a Zeiss Cytoscan is sent to a central PDP-10 for assessment. A large number of reports on investigations along these lines have been published [226-233, 238, 13-18, 20, 265]. To show that machine analysis could make discriminations difficult or impossible for humans observing standard cytologic preparations, Wied and his colleagues selected several difficult clinical problems: (1) discrimination among uterine glandular cancer cells, normal glandular cells and histiocytes [226, 229], (2) discrimination among uterine glandular cancer cells, cells from a possibly precancerous condition of the uterus (hyperplasia), and normal uterine glandular cells [232], (3) discrimination between similar-appearing normal cells from different parts of the uterus (endocervix and endometrium [231, 228],

(4) discrimination between normal and leukemic white blood cells (lymphocytes) [230, 15]. Also, studies were done in discrimination of cytochemically and morphologically identical tissue culture cells from human embryonic lung (HEL) and human epidermoid tumor (HEP) [227, 13, 14]. A variety of methodologies were tried. The usual histogram of optical density values received considerable massaging, being used both as a source of features (integrated or average optical density, sum of five highest density values with non-zero frequencies, sum of frequencies of specific density values) for constructing linear discriminant functions, and as a means of calculating threshold values for segmenting the gray-value image. In the latter technique, a composite optical density histogram derived from the entire training set population for one cell type is re-partitioned into unequal bin-widths so that each bin has approximately the same frequency of occurrence. (see Section VIB.)

This is conceptually the same as recoding the image using a maximum-entropy gray-level quantization, thresholding, and applying simple shape-recognition procedures to selected binary isodensity discriminants.

Wied and his colleagues have experimented with a large number of features and a wide variety of classification techniques. Recent development of system programming to automate the search for good features and classifiers should be a significant aid to exploration. However, published results do not appear to include investigation of the crucial question of performance of features and classifiers on images which were not contained in the training set used to design features and classifiers. This is typically the point at which pattern recognition systems have a tendency to fall apart. Also, the imagery used in Wied's studies was of a particularly high quality in terms of low noise, high resolution and accuracy. Whether Wied's methods can be used under less favorable

circumstances is questionable, although it should be pointed out that Wied's stated objectives do not include high-speed, fully automatic cell sample analysis.

2. Non-Morphological Analysis

Several systems have been developed using fluid-flow cell transport. These systems capitalize on the ability to control cell positions in a flow channel and typically make whole object (zero-resolution) non-morphological measurements.

The Rapid Cell Spectrophotometer (RCS) of Kamentsky/IBM [84-91, 94] is a general purpose instrument which has been applied to cervical cancer screening [94], blood cell differential counting [155], cell viability assay [87, 126] and exploratory studies to establish photometric features which could distinguish populations of functionally different cells [89, 73]. The RCS is capable of measuring 500 cells per second and has been equipped with a fluid-switch cell sorter.

In the cervical cell application, the RCS was subjected to a clinical trial in 1965 [94]. Performance for cervical swab specimens was not spectacular: 15% false negatives with 32% false positives, and for vaginal wash specimens 50% false negatives for 25% false positives. In addition, problems with sample preparation caused only 45% of the cervical swab and 70% of the vaginal wash specimens to be usable. This performance was apparently found to be unsatisfactory for clinical use.

The Coulter counter is a highly successful fluid transport system which has enjoyed extensive clinical and research use for establishing size distributions for a large variety of objects [253]. No optical measurements are made; the objects to be measured are pumped in an electrically conductive fluid medium through a small aperture in a non-conducting plate. A current is also flowing through the aperture.

Passage of an object causes a change in resistance through the aperture, generating an electric voltage pulse proportional to the volume of the object. This device was used [97] in an unsuccessful [68] experiment in cervical cancer screening. The Coulter counter has also been used as the front end of other fluid-flow systems, since the Coulter pulse can be used to control a variety of events.

An electronic cell sorter built by Fulwyler [54] combines a Coulter counter with an electrostatic droplet deflection system like the one used in the Inktronic printer, and can sort 1000 cells per second on the basis of cell volume. A stream of cell-containing fluid is broken into droplets by applying a vibration to the nozzle forming the stream. The Coulter volume signal is used to apply a proportional electrical charge to the droplet containing the cell. The droplets then pass through an electrostatic field which deflects the cells according to the volume-proportional charge.

The Automatic Multiparameter Analyzer for Cells (AMAC) proposed by Leif [105] is also built around a Coulter-effect synchronizer; having, however, a more elaborate sensor system with optical measurements available at several wave lengths. A system described by Hulett, e.al. [73] utilizes the droplet cell-sorter of Fulwyler, but substitutes a fluorescence measurement for the Coulter-effect volume signal used to charge the solute droplet.

A Crosland-Taylor laminar flow system by Van Dilla et.al. [211] uses a laser excitation source and a multichannel pulse-height analyzer to produce a frequency distribution histogram of fluorescent light emission per cell at a rate of up to 100,000 cells per minute. This instrument was used in a study to establish the time-course of DNA synthesis in a population of mouse fibroblasts.

These permutations and combinations of sensors, applications, and particle transport methods are summarized in Table 4.

Table 4. Flow Systems--Non-Morphological Measurements

Project/ References	Application	Measurement/ Performance	Transport/ Sorting Method
Fulwyler [54, 262]	Sort by volume while retaining cell viability to determine relation of volume to functional state	Coulter volume--Record data with multi-channel pulse-height analyzer. 30,000-60,000 cells/min. 96% viability	Electrostatic deflection of charged droplet system
Van Dilla [211]	Measure fluorescence due to DNA content in Chinese hamster ovary tissue-culture cells. Determine time of DNA synthesis and duration of cell phases. Improved statistical significance of large cell sample size over studies using scanning microspectrophotometers	Fluorescent emission (Feulgen) with Argon ion laser source at 488 nm. Record data with Multi-channel pulse-height analyzer. 10,000-100,000 cells/min.	Crosland-Taylor laminar-flow system. No sorting.
Rapid cell spectrophotometer [84-91, 94, 126, 155]	Cervical cancer screening. Live/dead cell assay. Differential white blood cell count. Population studies.	Scattering of white light. Absorption at various UV wavelengths. 60,000 cells/min.	Capillary tube transport with fluid switch sorting.
Mullaney [143-145]	Derive volume spectra of large cell populations. Increased accuracy over Coulter counter for fixed cells.	Volume measured by narrow-angle light scattering (primarily diffraction). Laser light source at 632.8 nm. Data stored in a multi-channel pulse-height analyzer. 10,000-100,000 cells/min.	Crosland-Taylor laminar-flow transport. No sorting.
Hulett [73]	Separate mixed cell populations on the basis of fluorochromasia developed by enzyme action on FDA. Cell viability is retained.	Mercury-arc excitation source. Light pulse from fluorescent cell activates charging pulse.	Electrostatic deflection of charged droplet stream (see Fulwyler).
Technicon Corp. Hemalog-D	Differential white blood cell count (Count white cells classified in five types).	Measures fluorescence and absorption at various wavelengths.	Capillary tube transport. No cell sorting.

IV. EQUIPMENT AND PROCESSES USED

This chapter describes some facilities which were used in this study. Figure 7 is a block diagram showing the interconnection of some of the hardware used to acquire and process the images. Digitized images from programmable film and microscope scanners were loaded into the Illiac III core memory under control of an interim software system, Show-and-Tell [171], which provides operator control of image selection. In addition, Show-and-Tell provides communication via a high-speed data link with an on-line image analysis package, PAX II, running on the IBM 360/75. PAX II is a parallel-image-processing language developed by the University of Maryland from the original simulator of the Illiac III Pattern Articulation Unit [80]. In the experiments in texture analysis, texture samples obtained interactively using Show-and-Tell were fed into programs developed by Jayaramamurthy, Tareski and Raulefs to implement the vari-valued logic approach to pattern recognition.

A. Programmable Scanners

A key aspect of this study is the use of a flexible programmable scanner in which the parameters of the image acquired are under computer control. The scanners used are the Illiac III's film scanner and microscope scanner. Both are of the flying-spot type in which the light source is a small (ca. 1 mil) spot of light generated by a Cathode Ray Tube. In the microscope scanner, light from the spot is projected down through the microscope ocular and is focused by the objective to a smaller spot on the specimen. A Photomultiplier tube (PMT) detects the light transmitted by the specimen. Another PMT independently examines the CRT spot brightness. These two PMT outputs are processed by analog

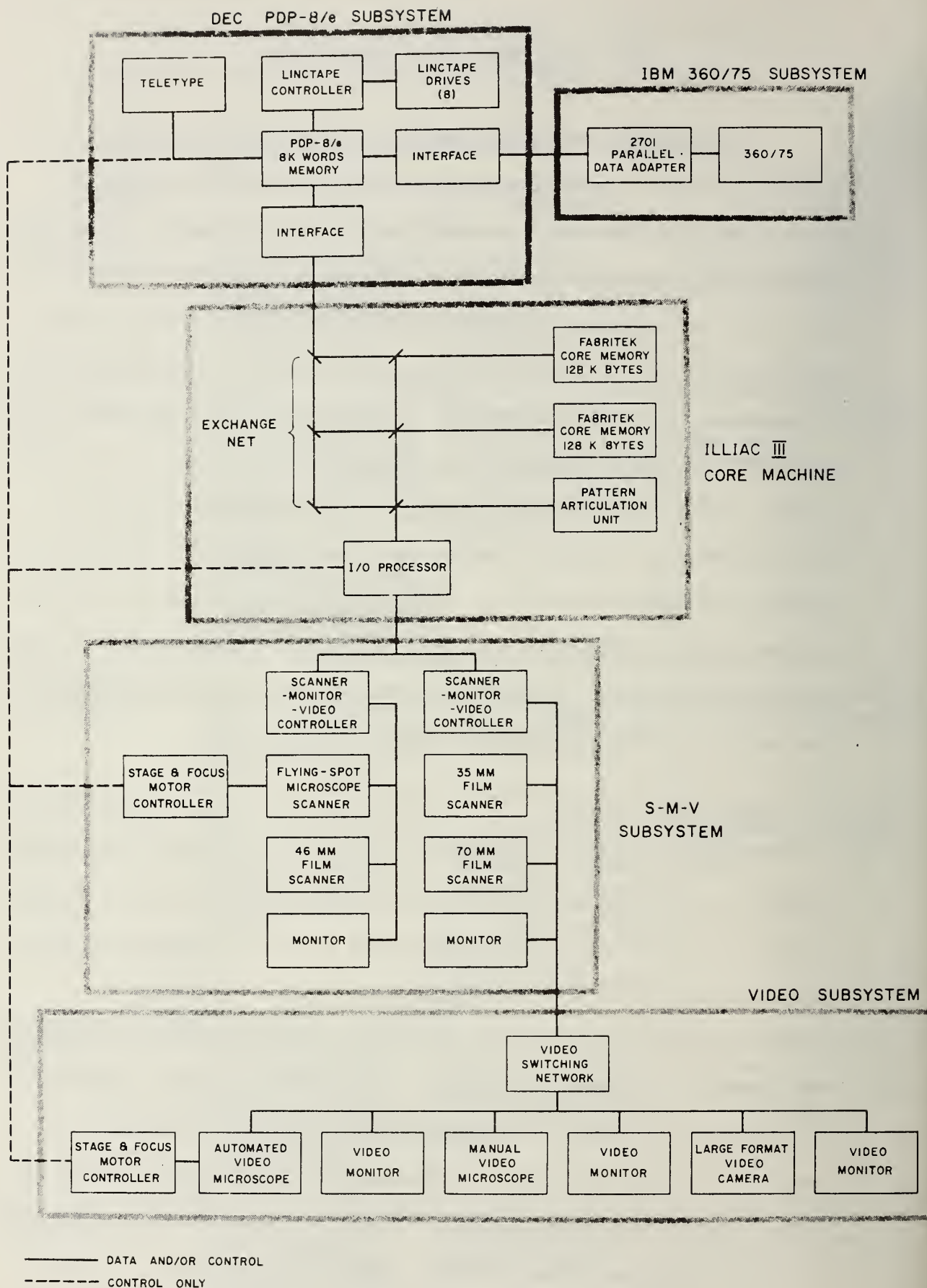


Figure 7. Block Diagram of Equipment Used

techniques to yield a signal whose amplitude is not significantly influenced by variations in the intensities of the CRT spot. This signal is digitized, currently to four bits, and the resulting numbers are loaded into Illiac III core at a rate of approximately 500,000 per second. By deflecting the spot in a regular, TV-like raster, a digital representation of the image is formed in the Illiac III memory. Following the notation of Rosenfeld [181], this representation will be indicated by (a_{ij}) , $1 \leq i \leq m$, $1 \leq j \leq n$, where m and n are the horizontal and vertical dimensions (in picture elements) of the conceptual array storing the picture. The resolution, proportions and placement of the sampling raster are under computer control. Up to a 256 x 256 element image representation can be contained in the memory as currently configured.

The film scanner works in a similar fashion, except that the raster and spot size conform to the format of a 46-mm. film frame.

A monitor system permits viewing the digitized images, either in real-time with the scanner, or by retrieving the picture element values from memory. The adjustment of display parameters (magnification, location on the screen) to improve the interpretability of the monitor display is also under computer control. Since the specimen or film is available on-line, it is not necessary to read in the entire image at one time. Instead, the image source can be treated as a read-only memory of very large capacity, and accessed only as needed. The flexibility of the scanner permits a quick coarse-scan look at large areas of the specimen followed by high-resolution interpretations of areas of interest.

The microscope stage motion and focus are also under control of the PDP8e/Show-and-Tell, via digital stepping motors.

B. Show-and-Tell

Show-and-Tell is an interactive programming system designed to permit on-line development and testing of scanning, preprocessing and feature extraction programs. The early version used in these experiments was intended to provide control of image acquisition, display, and real-time communication with an IBM 360, which would then provide image analysis via PAX II (see next section). [171]

In a typical development session, a programmer codes the processing program as an IBM 360 FORTRAN subroutine, with calls to PAX II as required. The subroutine can include calls to Show-and-Tell to type messages to the programmer, read data typed by the programmer, display or scan images, and transfer pictures to and from the 360. This program is submitted through the OS/360 batch system. When it begins execution, the programmer is informed, and he may begin testing his program by executing a Show-and-Tell CALL statement. Intermediate results can be displayed on the monitor, and various parameter values tested. Other commands can cause images to be saved on 360 tape or disk for later use as data for off-line testing.

C. Parallel Image Processing and PAX II

Over the past twenty years or so, a paradigm has evolved for reducing image data: a series of local operations converts the input image into an image or set of images from which the desired data can be relatively easily obtained. This approach apparently originated with Selfridge and Dineen [272, 273] and has been followed by many others, e.g. [258, 257, 270]. For a digital input picture, (a_{ij}) , a local

operation generates a new picture (b_{ij}), where the value of each output picture element, b_{ij} , depends on the picture elements in some relatively small neighborhood of each a_{ij} . Such an operation can be performed simultaneously for all the elements in a picture, since the output at each point depends only on the original values of the neighbors. This definition follows that of Rosenfeld, who has investigated the alternative case where each new value depends also on the new values of some of the neighbors [247]. In this case the local operation must be implemented serially. Rosenfeld showed that parallel local operations can in principle do anything serial local operations can do and vice versa, although with efficiency tradeoffs. If the local operations are homogeneous, i.e. the same function rule is applied at each location, then it becomes economically feasible to build a hardware processing array with a shared control to efficiently implement the operations. The hardware array processor can be expected to perform parallel local operations with a throughput improvement over conventional organization on the order of $pq:1$, where p and q are the dimensions of the array. This was in fact realized in the Pattern Articulation Unit of the Illiac III, which consists of 1,024 identical processors (stalactites) arranged in a 32×32 array [123]. Each processor can communicate directly with as many as eight (rectangular topology) or six (hexagonal topology) nearest neighbors. Local operations involving more distant neighbors can be performed by shifting. The fundamental instruction set, discussed in detail in the Illiac III system manuals [275], includes four basic classes of instructions: (1) forming logical functions of the contents of each stalactite and nearest neighbors (including simple arithmetic

functions); (2) shifting; (3) propagation; and (4) image loading, marking and readout of results.

These same instructions were simulated on a conventional computer (an IBM 7090) in the original PAX picture processing software system. By treating the accumulator as one planar row of the PAU, a modicum of parallelism was achieved. Workers at the University of Maryland extended and improved the PAX system, and embedded it in FORTRAN IV [80]. The resulting PAX II system therefore combines the numerical processing convenience of FORTRAN with a set of FORTRAN-callable subroutines to implement the parallel instructions. A data management system for storing and accessing the picture arrays is also provided.

D. Texture Recognition using Varivalued Logic

Varivalued logic, introduced by McCormick and Michalski [136, 137] is an extension and generalization of the binary-valued logic applied to switching theory. In this thesis, algorithms developed to support the theory of varivalued logic were used to automatically generate local operations capable of discriminating between two textures of cytological interest. As mentioned above, a local neighborhood of an element in a digital picture is some subset of the nearby picture elements. For example, in a rectangular sampling array each picture element and its nearest neighbors to the east, west, south, southeast, and southwest form a 3 x 2 local neighborhood. An $m \times n$ local neighborhood can be represented as an mn -dimensional vector, e.g.,

$$e = (x_1, x_2, \dots, x_{mn})$$

and can be regarded as an event in an mn -dimensional sample space. If the digital picture was quantized to h gray levels, then the sample space contains h^{mn} distinct events.

In the procedure described below, a training set of digitized samples of both textures is obtained. Using methods of statistical decision theory, each different event (local neighborhood) occurring in the training set is assigned to one or the other of the textures. Events in the sample space that did not occur in the training set are assigned to a DON'T CARE class. Having done this, one of the textures is regarded as a "true" set and an analogy can be drawn between the events assigned to the true class and the minterms of the disjunctive normal form of a switching function. This analogy is pursued and a suitably modified minimization procedure discovers an entity called an "interval cover," that defines a simplified categorizer preserving the assignment of events to texture classes much as a prime implicant cover defines a simplified switching circuit preserving some desired truth table.

The details of the procedure are presented by means of a simple one-dimensional "texture" example. The purpose is to communicate a general understanding of the concepts used in this thesis; a precise, formal exposition by the originators of the theory is contained in [136] and [137].

(1) Defining the categorizer by means of signal detection theory:

The purpose of this phase of the procedure is to assign each event in the training set to one or the other of the texture classes so that an unambiguous "true" set is defined for input to the interval covering process. Statistical decision theory provides a systematic way of doing this so that certain objectives are fulfilled optimally.

Let E^1 and E^0 be the sets of different events obtained from the texture samples T^1 and T^0 , respectively, and let

$n_1(e_k)$ Number of occurrences of the event e_k in T^1 (the number of "hits").

$n_0(e_k)$ Number of occurrences of the event e_k in T^0 (the number of "false alarms").

nT^1 . Number of events in T^1 .

nT^0 Number of events in T^0 .

In the one-dimensional example of Figure 8(a), the events are 1×3 local neighborhoods; the "textures" have been quantized to four levels, so the three-dimensional sample space contains 64 possible events. The nine events of E^1 and the seven events of E^0 are listed in the first columns of Table 5.

This information can be used to effect a disjoint partition of $E^1 \cup E^0$ that is optimal in the sense that certain decision objectives are satisfied as well as they can be, given the inherent separability or non-separability of the sample data. For example, it may be assumed that all misclassifications are equally costly and that one simply wants to minimize the number of errors. It can be easily demonstrated [59] that the sum of the probabilities of the two types of errors (saying that an event e is from T^1 when it is really from T^0 [a false alarm] and vice versa [a miss]) is minimized when an event is categorized in T^1 if its likelihood ratio (LR) is greater than 1.

The likelihood ratio of an event e_k is defined as

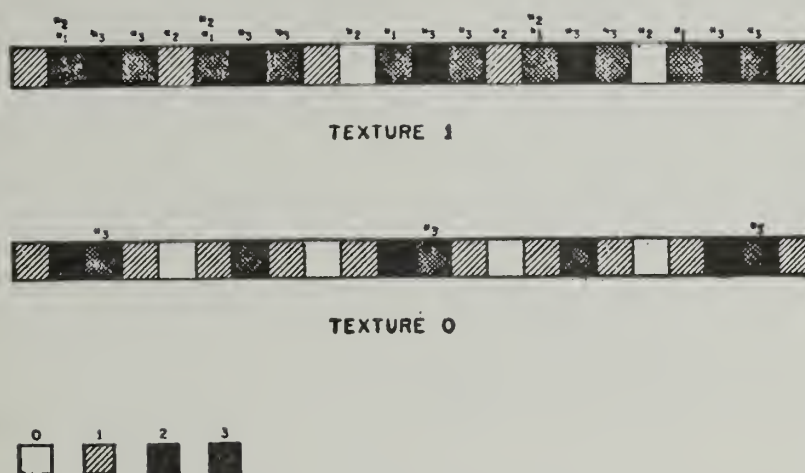


Figure 8(a). One-dimensional "Textures,"
Quantized to Four Gray Levels

Table 5. Statistics Derived from the One-dimensional Texture
of Figure 8(a) using a 1×3 Template to Define Events

Event (e)		$n_1^{T_1}$ (e)	$p(e T_1)$	$n_0^{T_0}$ (e)	$p(e T_0)$	$LR(e)$
	(1,2,3)	3	.15	0	0	-
	(2,3,2)	5	.25	0	0	-
	(2,1,2)	2	.10	0	0	-
	(1,0,2)	1	.05	0	0	-
	(0,2,3)	2	.10	0	0	-
	(3,2,0)	1	.05	0	0	-
	(2,0,2)	1	.05	0	0	-
	(3,2,1)	4	.20	3	.15	1.33
	(2,1,0)	1	.05	4	.20	.25
	(1,3,2)	0	0	3	.15	0
	(1,0,1)	0	0	4	.20	0
	(0,1,2)	0	0	2	.10	0
	(1,2,1)	0	0	2	.10	0
	(0,1,3)	0	0	2	.10	0
	(0,1,3)	0	0	2	.10	0

$$LR(e_k) = \frac{P(e_k|T^1)}{P(e_k|T^0)}$$

where

$$\begin{aligned} P(e_k|T^1) & \text{ The probability of } e_k \text{ occurring in } T^1, \text{ i.e., the proba-} \\ & \text{ bility of } e_k \text{ conditional on } T^1, \\ & = n_1(e_k)/nT^1. \end{aligned}$$

$P(e_k|T^0)$ is defined similarly.

It can be seen by looking at Table 5 that this decision goal is realized in the one-dimensional example. If only those events with $LR > 1$ are called T^1 , then four misclassifications (three false alarms and one miss) result. If any other subset is classed T^1 , more errors occur. Other decision goals can be realized by using a different likelihood ratio threshold [59]. For generality the threshold can just be designated β .

With this background, we are ready to partially define the categorizer Ψ_R on the basis of the training set information.

Let

E = Event space.

$$F^{1\beta} = \{e | e \in E^1 \cup E^0 \text{ and } LR(e) > \beta\}.$$

$$F^{0\beta} = \{e | e \in E^1 \cup E^0 \text{ and } LR(e) \leq \beta\}.$$

$$F^* = \{e | e \in E \setminus (E^1 \cup E^0)\}.$$

Then define Ψ_R by its acceptance set R , i.e.,

$$\Psi_R(e) = 1, \quad \text{iff } e \in R$$

where

$$F^1 \beta \subseteq R \subseteq F^1 \beta \cup F^*$$

and

$$R \cap F^0 \beta = \phi.$$

Note that the determination of which events in F^* are in R has not been made at this point; these represent DON'T CARE events that are assigned as described in (2) below.

The receiver-operating-characteristic (ROC) curve is a useful device for observing and predicting the behavior of these categorizers. To make the curve, each event $e \in E^1 \cup E^0$ is regarded as a two-component vector with $x = p(e|T^0)$ and $y = p(e|T^1)$. An ordering can be imposed on these vectors by sorting them in descending order by the likelihood ratios of the e . The curve is generated by placing the tail of the first vector at the origin and then concatenating the rest in order. For the one-dimensional example, the graph shown in Figure 8(b) is the result. The ROC displays several useful items of information in an easy-to-see form. For one thing, the training-set performance of a categorizer for each value of β is shown directly, since for each threshold the y coordinate is equal to

$$\sum_{\{e|LR(e)>\beta\}} p(e|T^1)$$

and the x coordinate is equal to

$$\sum_{\{e|LR(e)>\beta\}} p(e|T^0).$$

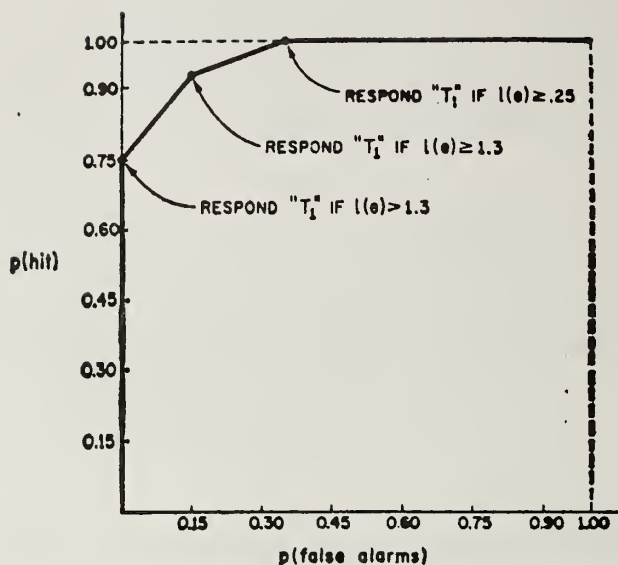


Figure 8(b). Receiver Operating Characteristic
Corresponding to Table 5

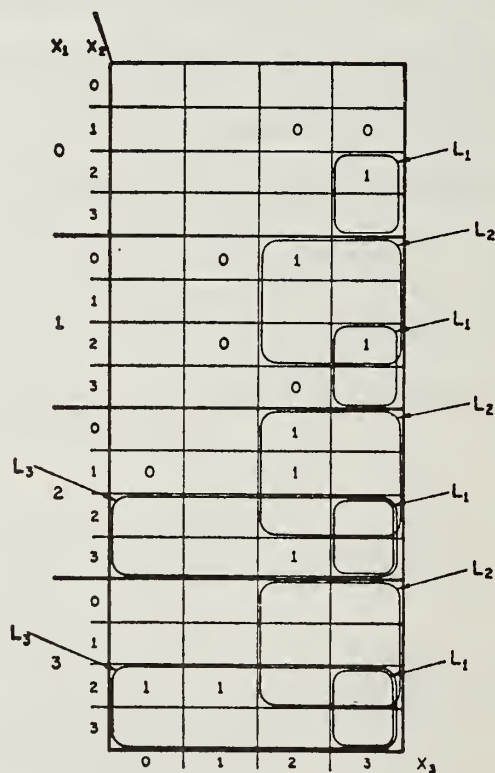


Figure 8(c). Generalized Logic Diagram with Interval Covering
of T^1 against T . $L_1 = X_2 X_3$; $L_2 = X_1 X_2 X_3$; $L_3 = X_1 X_2$

The point on the ROC corresponding to a given value of β is easy to find since it is the tail of the vector with slope = β . (Note that β only has a finite number of values with different performance effects.) The ROC also provides a measure of the inherent separability of the textures in the training set. The area under the curve is equal to 0.5 if the textures are nondistinguishable (all events occur with equal probability in both textures) and is equal to 1.0 if the textures are perfectly distinguishable (all events occur in one or the other texture but not both).

(2) Implementing the categorizer by means of varivalued logic:

In principle the local categorizer described in the preceding section could be implemented by just looking up input events in a table of events and likelihood ratios. However, for real textures and useful neighborhood sizes this process would be hopelessly slow. Also, no categorization would be performed for events not in the training set. By applying some concepts from switching-theory, equivalent but much more efficient categorizers can be generated. This is accomplished by a technique analogous to switching-theoretic procedures for minimization of the disjunctive normal form of a switching function.

If Table 5 is viewed as a truth table where events in $F^{1\beta}$ are true and the others false, then the disjunctive normal form can be expressed as $V\xi_i(e_i)$ where ξ_i is a predicate that has output true when the input is a particular event e_i from $F^{1\beta}$, and output false otherwise. The symbol V represents the logical OR of the predicates. McCormick and Michalski have developed a generalization of switching theory [137] that permits the transplantation of much of the minimization machinery already in existence. In particular, Michalski's A^Q algorithm for generation of quasi-minimal

covers can be used. To explain the method it is necessary to introduce a few items of notation from [137].

E is the event space as before, i.e., the set of all events

$$e = (x_1, x_2, \dots, x_n), \quad 0 \leq x_i \leq h-1.$$

A literal, ${}^{a_i}_{X^i} b_i$ is the set of all events $e \in E$ whose i -th component lies between a_i and b_i .

$${}^{a_i}_{X^i} b_i = \{(x_1, x_2, \dots, x_n) \mid a_i \leq x_i \leq b_i\}.$$

An interval is a set-theoretic product of literals

$$L = \bigcap_{i \in I} {}^{a_i}_{X^i} b_i \quad I \subseteq \{1, 2, \dots, n\}.$$

The interval represents a "box" in hyperspace that includes all events between (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) . Note that components not specified by the interval are free to take on any integer value in $[0, h-1]$.

An interval cover of the set $F^{1\beta}$ against $F^{0\beta}$ is defined as a union of intervals L_j such that

$$F^{1\beta} \subseteq \bigcup L_j \subseteq F^{1\beta} \cup F^*.$$

Thus an interval cover contains all the events in $F^{1\beta}$ plus some in F^* , but none in $F^{0\beta}$. However, the interval cover will represent this partitioning of the space of possible events much more concisely than just enumerating all the events that should be labeled T^1 . Also, the interval cover can classify events that were not in the training set because of

the inclusion of F^* events in the boxes. A quasiminimal cover can be generated via the A^Q algorithm which can only be sketched here.

Referring again to our usual example, an interval cover (a minimal one, as it turns out) can be generated manually by means of a visual aid, the generalized logic diagram (GLD), which was introduced by Michalski [136], Figure 8(c). In the particular case when variables assume only two values, the GLD reduces to a diagram that resembles a Marquand-Veitch diagram. The GLD is a representation of the entire event space, $E:64$ events in this case. To use it, the events of $F^{1\beta}$ are mapped in as ones and those of $F^{0\beta}$ as zeros. The squares left over represent F^* (DON'T CARES). The cover is found by an iterative procedure that begins by picking the first "one" encountered in a TV-like scan of the GLD and discovering all of the maximal intervals that include that one, but no zeros (an interval "star"). One of these, the interval including the most ones, is added to the covering set (initially empty). All of the ones included in the star are temporarily eliminated and the scan of the GLD is resumed. The first one encountered is again selected and the iteration repeats. Eventually all the ones have been eliminated. If all ones are included in the covering set, then the cover is minimal. Otherwise, the cover is patched up to include the neglected events, and may not be minimal.

This procedure was followed for the example and a minimal covering using three intervals resulted (Figure 8 (c)).

$$L_1 = {}^2x_2^3 {}^3x_3^3, \quad L_2 = {}^1x_1^3 {}^0x_2^2 {}^2x_3^3, \quad L_3 = {}^2x_1^3 {}^2x_2^3.$$

If these are used to form a categorizer Ψ_R , where $\Psi_R(e) = 1$ for $e \in R = L_1 \cup L_2 \cup L_3$, then the event categorization shown in Figure 8 (a) results. The asterisks appearing above Texture 1 and Texture 0 indicate events for which Ψ_R had output equal to 1. The subscripts on the asterisks denote the interval producing the hit. Notice that the probability of a hit ($19/20 = 0.95$) and the probability of a false alarm ($3/20 = 0.15$) is as predicted by the ROC for the likelihood ratio decision rule with $\beta = 1$. These textures could easily be discriminated by labeling regions with hit density over some averaging aperture greater than, say, 55 percent as Texture 1. In a digital parallel processor, like the Illiac III's pattern articulation unit, Ψ can be implemented as a local operation. The input to the operation is a digitized picture in several gray values and the output is a binary plane labeling each element in the input picture as to which texture the picture element most likely belongs. The application of simple smoothing or noise-removal algorithms would then make segmentation into texture regions relatively easy.

V. CYTOLOGICAL IMAGE PROCESSING EXPERIMENTS

A. Blob Detection

One cell component which must always be present in a cell is the nucleus. An object on the slide is not a cell unless it has a nucleus, or at least if it has no nucleus it is of no interest for present purposes. In stained preparations, the nucleus is also the most invariant of the cell components in the sense that it is almost always a fairly symmetrical, round object which is darker than its surroundings. It makes sense, then, to begin the smear analysis process by testing the image for the presence of these round dark objects, which are referred to below as blobs. As has been discussed in Chapter III.C.1, it is necessary that blobs caused by clumps of leukocytes be masked off in some way. In addition, the blob finder should operate on images of low resolution which contain some noise.

A blob detector with these characteristics was implemented in the form of the PAX II subroutine, 'BLOB'. A listing of BLOB (and the other subroutines mentioned in this thesis) is provided in the Appendix for readers with knowledge of PAX II. BLOB is essentially a predicate which has output TRUE wherever a region satisfying the parameters is sensed. BLOB functions by examining the neighbors of each picture element, a_{ij} , looking for a decrease in gray value occurring at a distance falling in the range $[IRADLB, IRADUB]$, where IRADUB and IRADLB are parameters passed to the subroutine. BLOB requires that all neighbors at a distance less than IRADLB be no more than IDROP less than a_{ij} in gray value, where IDROP is another parameter. On the other hand, at least one neighbor which falls in the distance range $[IRDLB, IRADUB]$ along each principal direction must be more

than IDROP less than a_{ij} . Thus, BLOB looks for points falling on local hillocks in the gray-value topography of the image by sending a probe in each of eight directions from each a_{ij} , and requiring that the gray value drop off before distance IRADUB, but that the gray value not drop off before distance IRADLB. Further, a "hit-count" is maintained for each a_{ij} of the number of such radii which found drop-offs as specified. If this number exceeds INTOL (another parameter), then the output plane is set to 1 at location (i,j) . Figure 10(a) shows the operation of the intermediate filter which transforms the input gray-scale picture into a "hit-count" picture. The PAU architecture permits this algorithm to be executed simultaneously for all of the elements of the picture. The approach is similar in spirit to that used in some character recognition devices, such as the IBM 1975 [67] in which a loose-fitting template samples high-information parts of the pattern so that minor variations in type font and print quality are accommodated.

The subroutine BLOB is used in various ways in attacking the Pap smear problem. First, BLOB is applied to the picture with parameters such that leukocyte-sized blobs are detected. BLOB is capable of identifying individual leukocytes in images of fairly coarse resolution, even when they lie close to each other. Each picture element in the images used covers an area nominally $1.42 \times 1.95 \mu\text{m}$; the entire image covers a rectangle approximately $182 \times 250 \mu\text{m}$ on the specimen. The BLOB subroutine output plane ($128 \times 128 \times 1$ bit) contains a bit on for each picture element found to represent a small blob. After the small blobs have been located, BLOB is iteratively reapplied to the same input picture, this time with parameters defining blobs in the size ranges corresponding to various types of atypical cell nuclei.

This second group of output planes is logically ORed together, forming a composite "big blob" location plane. Using a propagation process which can be implemented by the PAU operation CONNECT, the small blobs are "smeared", so that each on bit is replaced by a small cluster of on bits [150, 270]. The smeared plane is inverted and ANDed with the big blob plane, thereby wiping out big blob locations near small blobs. In a big blob caused by a cluster of little blobs, it is assumed that at least one little blob location will be within smear range of the center of the big blob.

Given the BLOB subroutine described above, the primary problem is to determine the values which should be assigned to the parameters IRADUB, IRADLB, INTOL, IDROP (for big and small blobs) and the smear distance. In the big blob case, this was accomplished by making use of Table 2, which was derived from quantitative studies by Wied and Reagen discussed in Chapter II.C.1. Table 2 shows mean nuclear area for five classes of cancer-related conditions. If the nuclei are assumed to be nearly round, the area measurements are easily converted to approximate radius measurements by means of the relation $RADIUS = \sqrt{AREA/3.1}$. These can in turn be easily converted to digital raster distance, where distance between two picture elements is defined as the number of unit shifts required to carry one picture element onto another. A unit shift carries an element onto one of its eight (rectangular topology) nearest neighbors. Clearly, shifts in a diagonal direction cover more ground on the original image than do shifts parallel with the X or Y axes. In addition, the scanner system used here generates a raster with an aspect ratio nominally 8.5:11, so that a shift in the direction of the scan (Y) axis covers more ground on the original image than do shifts along the line

(X) axis. Measurements on scanned photomicrographs showed that the aspect ratio was actually closer to 8.0:11, so that picture element centers on a 128 x 128 scan of a 102 x photomicrograph were 1.95 μm distant from each other (on the specimen) in the Y-direction and 1.42 μm distant in the X-direction. In order to prevent these factors from introducing an anisotropism in the response of the blob detector, a correction was applied to the parameters IRADLB and IRADUB, which are assumed to be given in terms of the X-direction. The correction for diagonal distance was $1.42/\text{SQRT}(1.42^2 + 1.95^2) = .59$, and for vertical (Y-direction) distance was equal to $1.42/1.95 = .73$.

Three different raster distance values were established.

on the basis of information in Table 2:

<u>Condition</u>	<u>Cell Nucleus Radius (μm)</u>	<u>Raster Distance (X-Direction)</u>
Dysplasia	7.4	5
Ca <u>in situ</u>	6.0	4
Kerat. Ca	5.0	4
Non-kerat. Ca	5.3	4
Small-cell Ca	4.	3

where the X-direction raster distance is obtained by dividing the radius in μm by 1.42, rounding, and taking the integer part. Subroutine BLOB was applied three times to the input image shown in Figure 9(b), with parameters set each time to detect one of the size ranges corresponding to an atypical cell class of Table 2. INTOL and IDROP were set to eight and three, respectively, for all iterations. When the three output planes were ORed together, 129 hits resulted of which 82 were judged (by looking

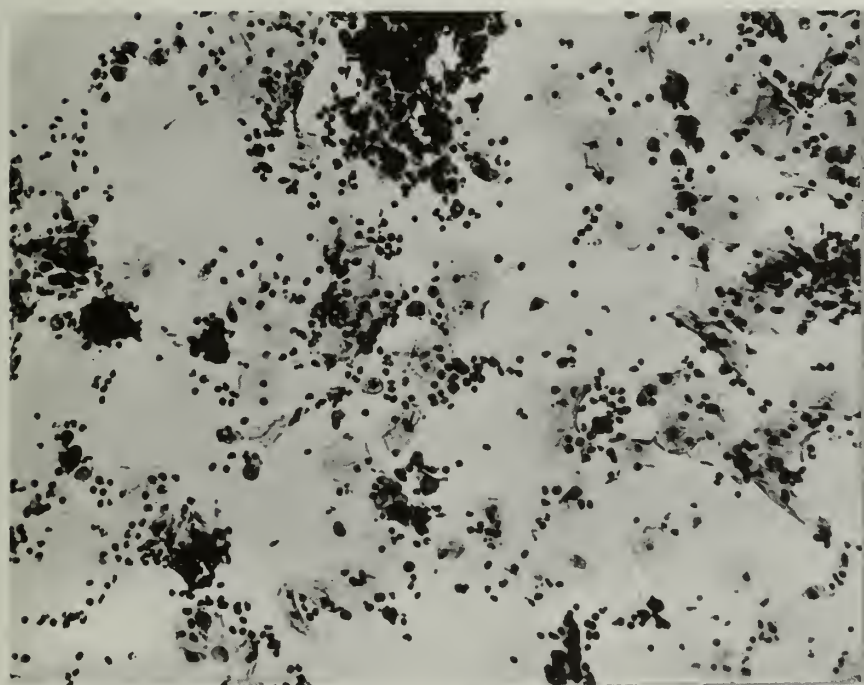


Figure 9(a). Photographic Image: Blobs

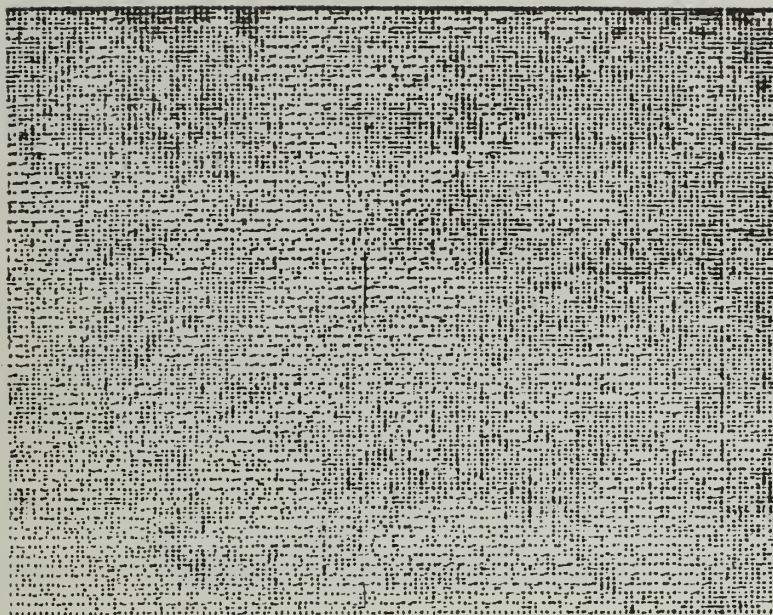


Figure 9(b). Digitized Image: Blobs

at the original photomicrograph) to be clumps of white cells and 47 were judged not to be clumps of white cells. Figure 11(a) shows these 129 hits circled. IRADLB and IRADUB were set equal to one for the clump mask, since at the low resolution, only one picture element per leukocyte is at maximum density. Values for INTOL and IDROP were determined experimentally by varying INTOL and IDROP and observing the overall error rate. Table 6 shows performance of the masking process on the training picture as a function of INTOL and IDROP, where performance is measured by counting the number of big-blob-detector hits which were visibly (in the original photomicrograph) caused by clumps of leukocytes and noting the rate at which these are successfully rejected by the masking procedure. An effort was made to eliminate effects of context on this subjective decision by viewing the objects through an aperture. A "False alarm" is a failure to mask off a clump, while a "miss" is the masking of a non-clump. Note that the major consideration in this experiment was the false alarm rate caused by leukocyte clumping. Thus, response by the blob detector when presented with other objects was not pursued, i.e., the big blob parameters and the consequent hit rate were assumed to be given, and the problem was defined as minimization of the overall error rate by adjusting the clump mask parameters. As indicated by Table 6, the optimal values of INTOL and IDROP were found to be INTOL=7 and IDROP=1 for this training image, resulting in a false alarm probability of .07, a miss probability of .66, and an overall error probability of .29. Figure 11(b) shows the effect of masking big blob hits using the optimal mask parameter values.

Table 6. Blob-detector Performance as a Function of
Small-blob Parameter Values

(Clumps of leukocytes are considered noise, other blobs larger than individual leukocytes are considered signal.) Figures 11(a)-(b) display processing using the optimal values of $\text{INTOL} = 7$, $\text{IDROP} = 1$.

Parameters		Performance			
INTOL	IDROP	FALSE ALARMS	MISSES	TOTAL ERRORS	PERCENT ERRORS
5	1	0	47	47	36
5	2	3	42	45	35
5	3	17	32	49	38
6	1	0	42	42	33
6	2	13	31	44	34
6	3	33	22	55	43
7	1	6	31	37	29
7	2	29	15	44	34
7	3	52	2	54	42
8	1	38	25	63	49
8	2	63	2	65	50
8	3	69	0	69	54

B. Cell Detection

As described in Chapter II, images of superficial cells can become extremely complex. It is therefore necessary to determine precisely the data to be extracted so that unnecessary processing is avoided. In the present work, it was decided to try to locate fairly well-preserved superficial cells which might be touching or overlapping each other, but not badly crumpled or contained in thick cell masses. The cells to be detected would be the same ones a human microscopist could readily count.

It was felt that it should not be necessary to perform detailed image segmentation to accomplish this, and in fact, that this would be impossible to do in the low-resolution images contemplated. Other workers [177, 106] have described algorithms to handle touching or overlapping cells, but usually postulate an idealized input image in which edges are unambiguously determined and noise-free. As can be seen in Figure 9, this is far removed from reality, and the algorithms would be useless if applied to the output of any conceivable preprocessing scheme operating on real images in real time. Precise delineation of edges is not required for the present purpose since no measurements are taken. The presence of a cell in a digitized image is established by applying a crude structural model of a flattened epithelial cell and reporting a hit where there is sufficient match, an approach which is similar to that taken in subroutine BLOB. In this case, the model must be somewhat more complex so as to account for the greater range of possible configurations of the complete cells.

Processing is based on the model of the cell image described previously in Chapter II.B.2, in which the cells of interest are composed of a region of fairly consistent optical density (the cytoplasm) with a darker blob (the nucleus) more or less in the center, as in Figure 3. Cell-detection proceeds by attempting to find blobs which are approximately centered between pairs of "step-down" edges. A step-down edge is an edge oriented orthogonal to a line radiating from a blob point, where the gray-value changes in the negative direction as the line crosses the edge. In principle, this process is insensitive to overlapping of cytoplasms, since the requirement for equally-distant negative-going edges will associate edge pairs with the correct nucleus, unless either (1) the cells are nearly coincident, in which case the distance between the nuclei may be less than the tolerance established for the equal-distance criterion or (2) the cells have off-center nuclei. In case (2), disambiguation may still be possible where the nucleus is centered in at least one direction.

The present implementation of this procedure takes the form of five PAX II subroutines: AVG, DIRDIF, RIDGE, RADAR and BLOB. AVG performs a digital low-pass filtering operation which is intended to decrease the effect of small detail and noise on the edge computation. AVG uses an efficient parallel algorithm described by Rosenfeld and Thurston [182], in which the gray value of each picture element is replaced by the average gray-value in an $n \times n$ square neighborhood of the element, when n is a power of 2. AVG does this computation in parallel using arithmetic of $n*2+4$ bits precision. The optimal size of n is a function of

the "conspicuousness" (Rosenfeld's term) of the edge to be detected. For this application, n was established by trial-and-error using the interactive facilities of Show-and-Tell. Subroutines DIRDIF and RIDGE cooperate to reduce the blurred image to a binary plane with ones marking the location of edges orthogonal to a given direction. DIRDIF performs a directional differencing operation on the AVG output by subtracting it from a shifted copy, where the shift is in the direction of interest and has magnitude n (the size of the side of the averaging area). RIDGE detects ridges (local maxima occurring in chains) in the difference picture and eliminates spurious maxima not aligned orthogonally to the given direction. The sign of the difference is available for use in establishing whether an edge is negative-going with respect to the given direction.

Subroutine BLOB is pressed into service again to detect the leukocyte-sized cell nuclei used as reference points for the cell-finding operation. In order to determine which of the blob points is centered between negative-going edges, another subroutine, RADAR, is used. RADAR causes each blob point to become a sort of Radar transmitter. The edge points for one sign and direction are the targets. Using CONNECT, the propagation operation of the PAU, a spreading beam of ones is propagated in a given direction (see Figures 10(b) and 12(e)). Edge points caught in the beam are identified by ANDing the beam with the edge point plane at each stage of propagation of the beam. These edge points are caused to reflect a signal back to the blob "transmitter" along the beam, again by using CONNECT. A plane containing the blob points which received return signals at each propagation stage is stored. RADAR is then re-applied in the opposite direction with edges of opposite sign as targets. The output, then, of RADAR applied in a pair of directions is two stacks of planes containing coded range information. Distance is represented by the position of the plane in the stack, and on bits in a

plane indicate which blob points received echoes at that distance. If successive pairs of planes, one from each stack, are ANDed, and the result is a plane having a bit on for each blob found to be approximately centered between a pair of step-down edges, the number of edge-pairs discovered for each blob can be used as a measure of the degree of "ill-formedness" of the cells. Cells detected at an ill-formedness index of two are shown in Figure 12(f).

C. Texture Processing

Drying artifact can cause large, dark blobs in cervical smears, which must be distinguished from cell nuclei to avoid excessive false positive reports. Drying artifact occurs when cells are not fixed promptly upon being spread on the slide and looks like the large masses featured in Figure 13(a). Chromatin is the result of staining the genetic material in the cell nucleus. When the chromatin shows a texture consisting of

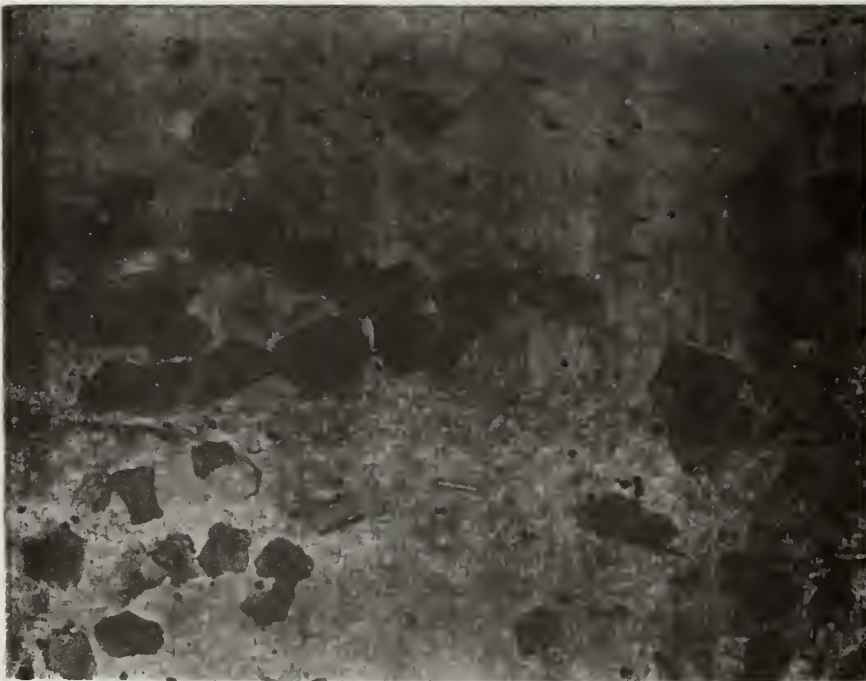


Figure 9(c). Photographic Image: Cells

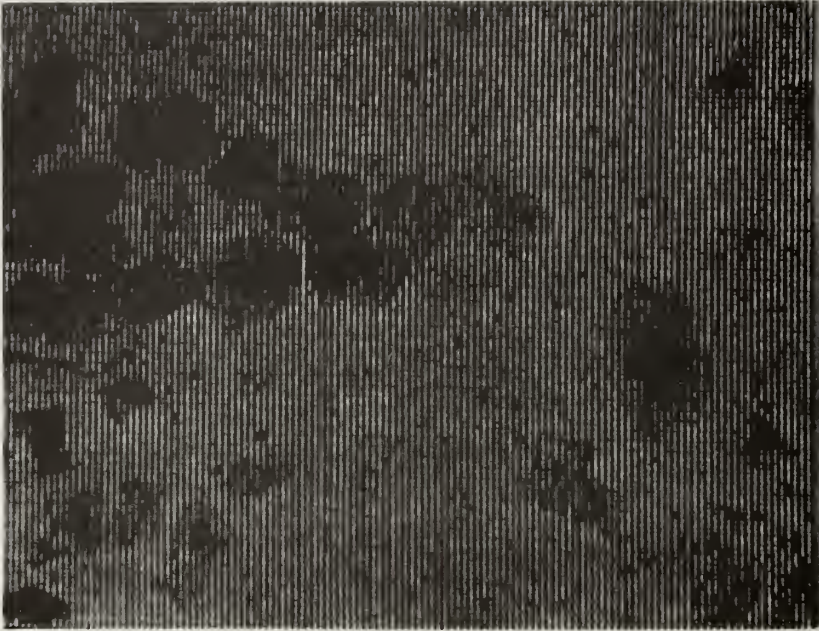


Figure 9(d). Digitized Image: Cells

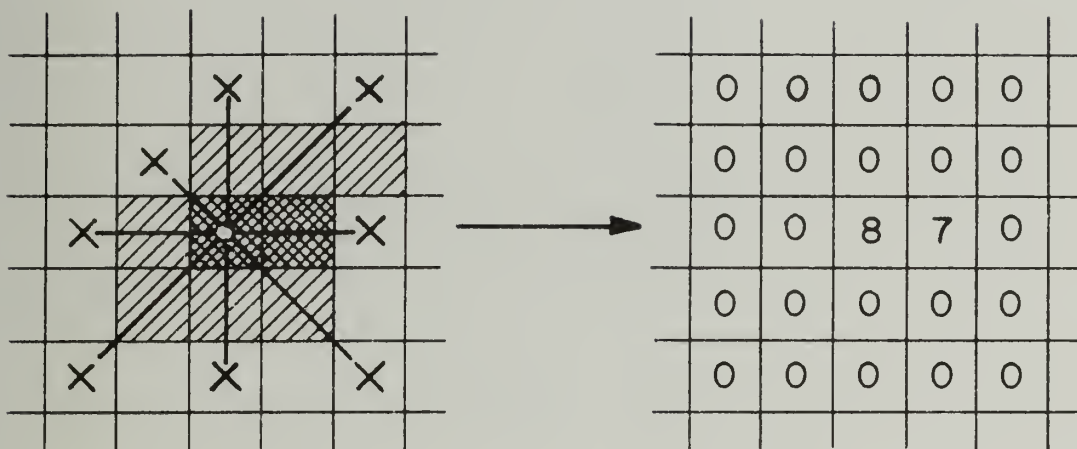


Figure 10(a). "BLOB" Operation. Three Gray Levels.

IRADUB = 2, IRADLB = 1, IDROP = 2

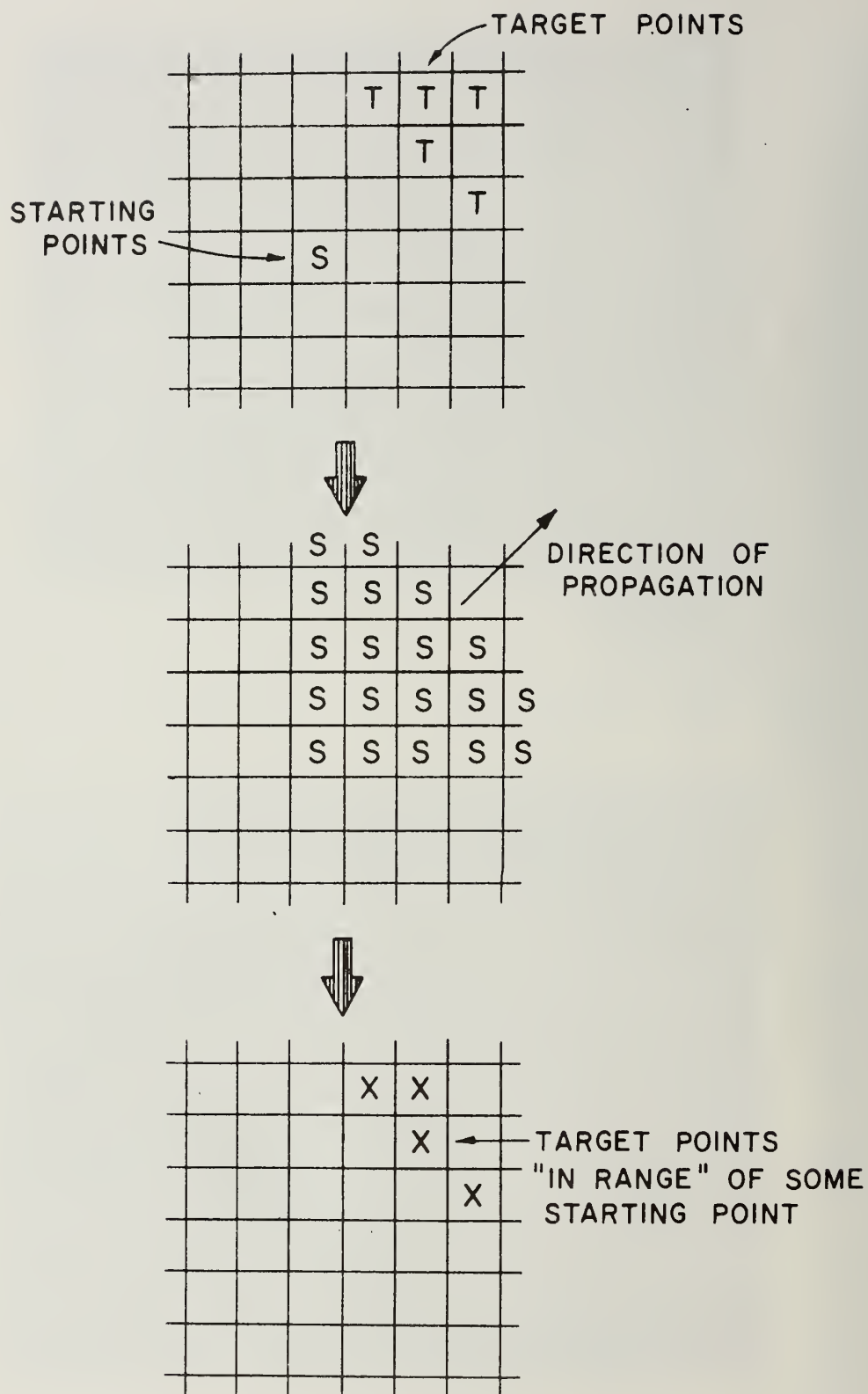


Figure 10(b). "RADAR" Operation

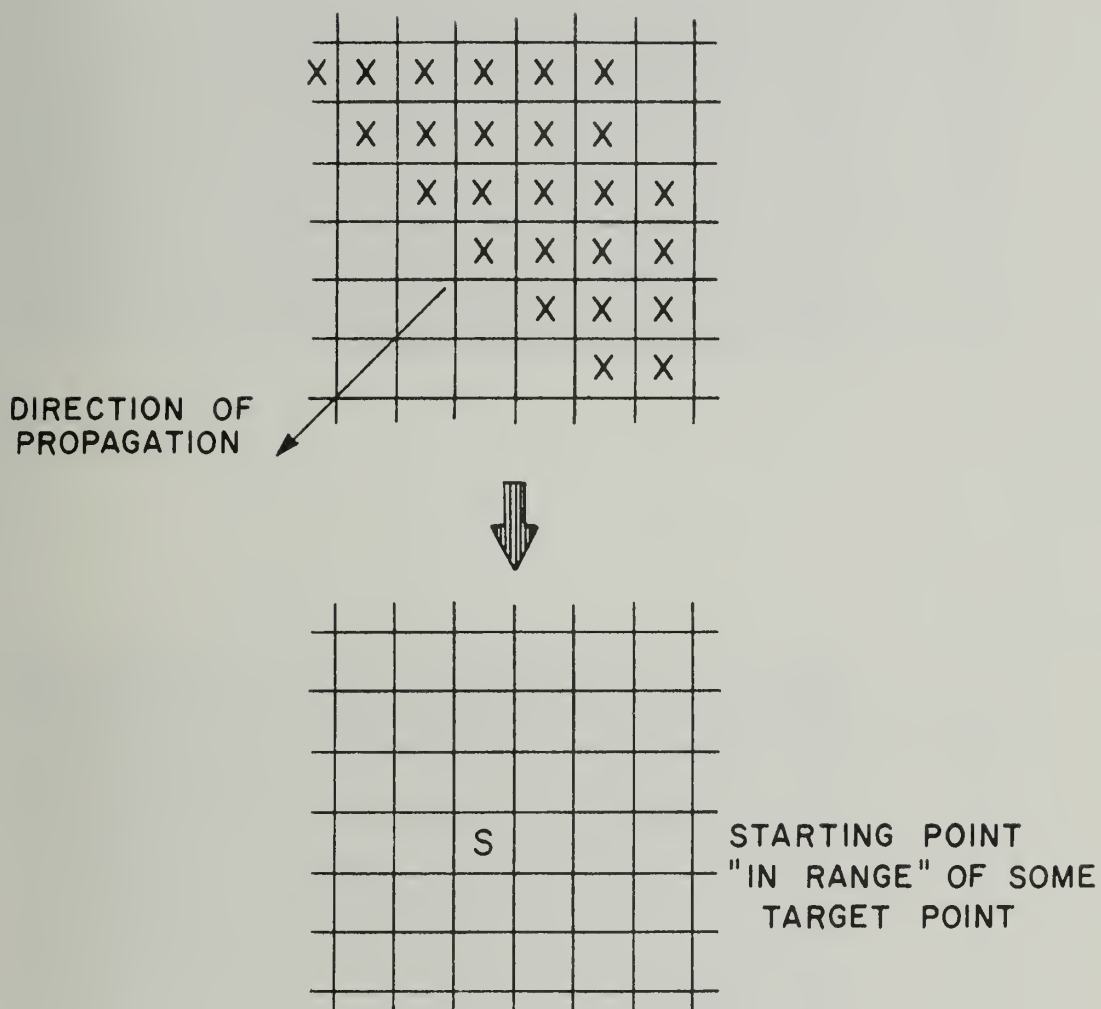


Figure 10(b) (continued). "RADAR" Operation

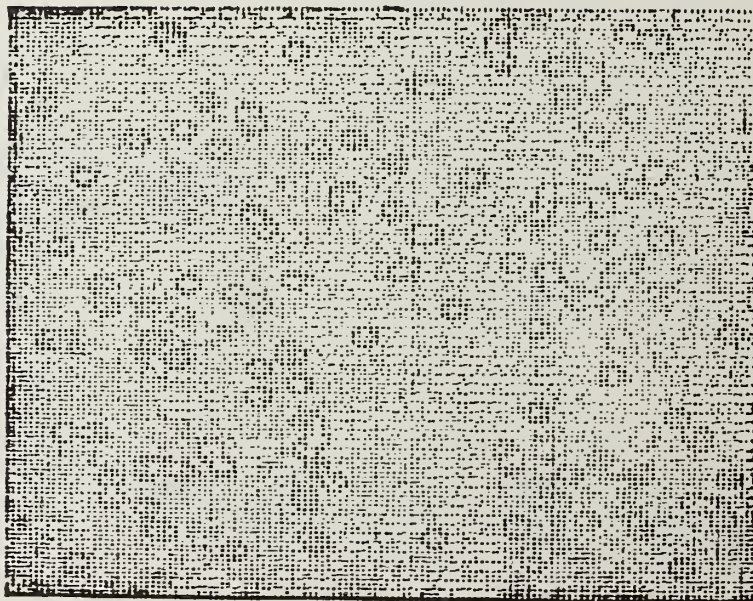


Figure 11(a). Large Blobs Detected

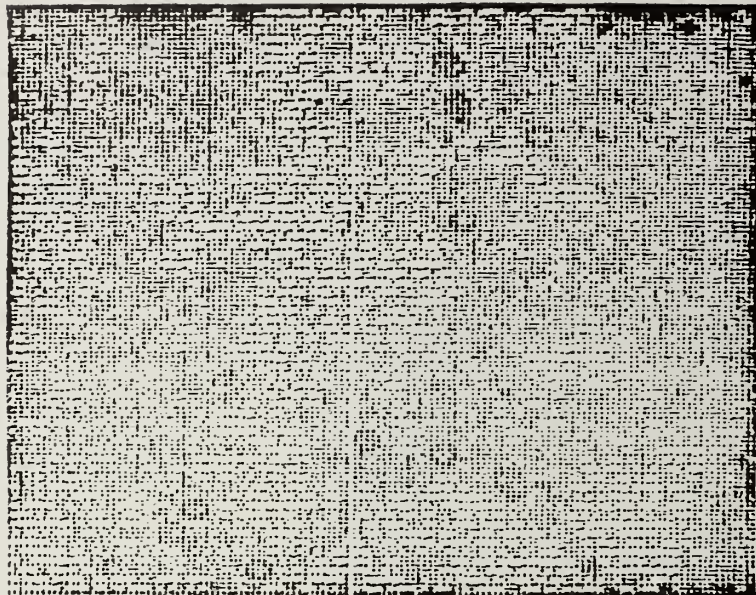


Figure 11(b). Blob Detector Output
(Clumps suppressed)

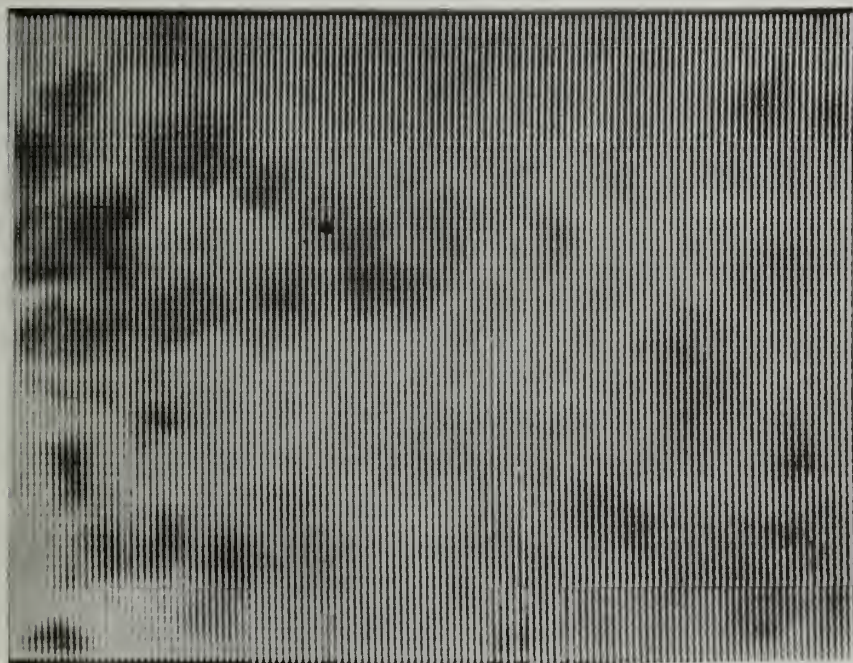


Figure 12(a). Smoothed Image (KPOW = 2)

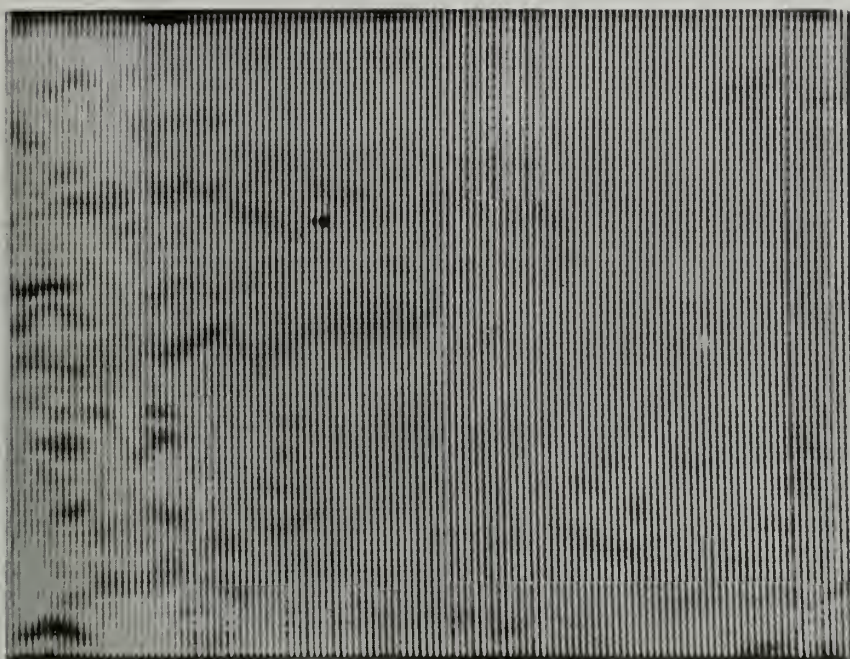


Figure 12(b). Magnitude of Differences in Y-Direction

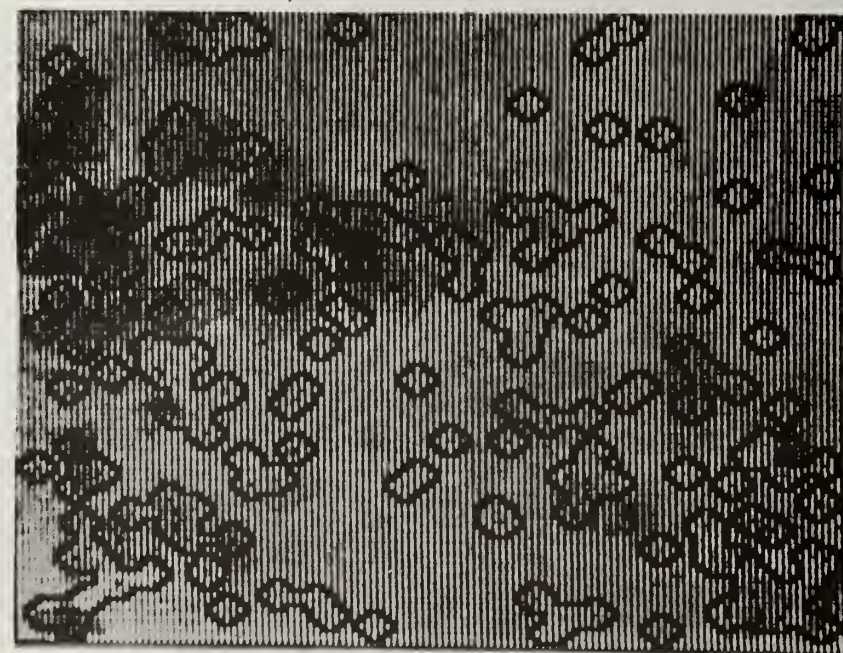


Figure 12(c). Blobs Detected

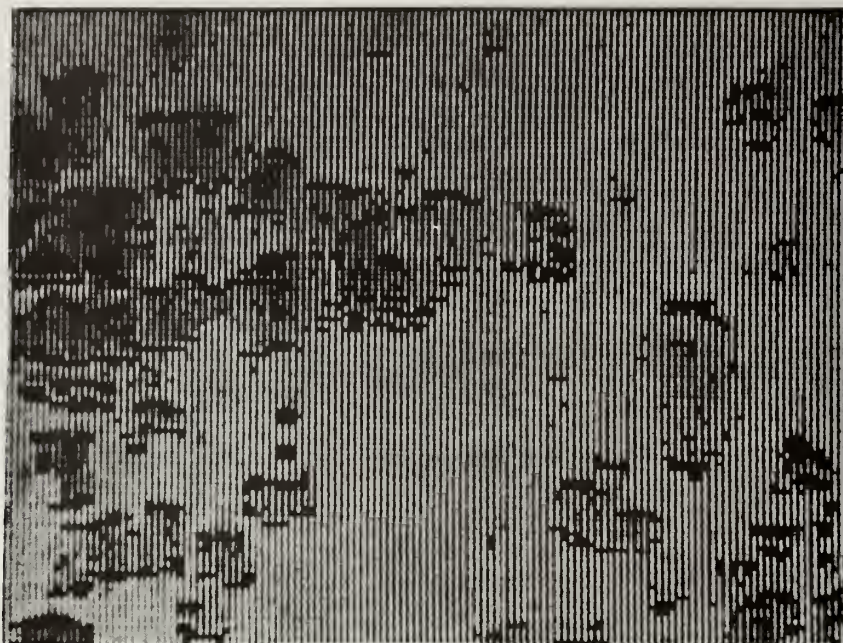


Figure 12(d). Cytoplasmic Borders
in the X-Direction Detected

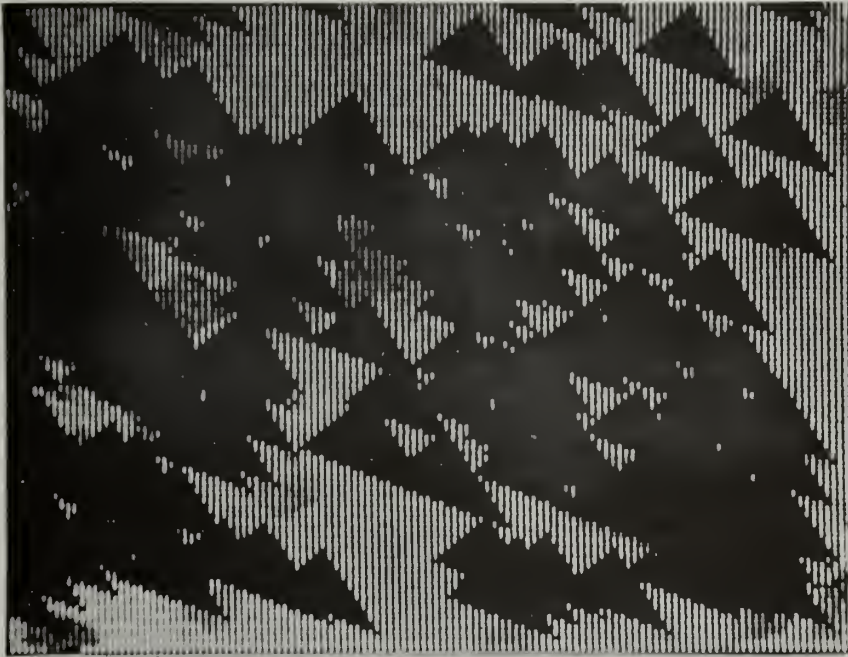


Figure 12(e). RADAR "Beams"

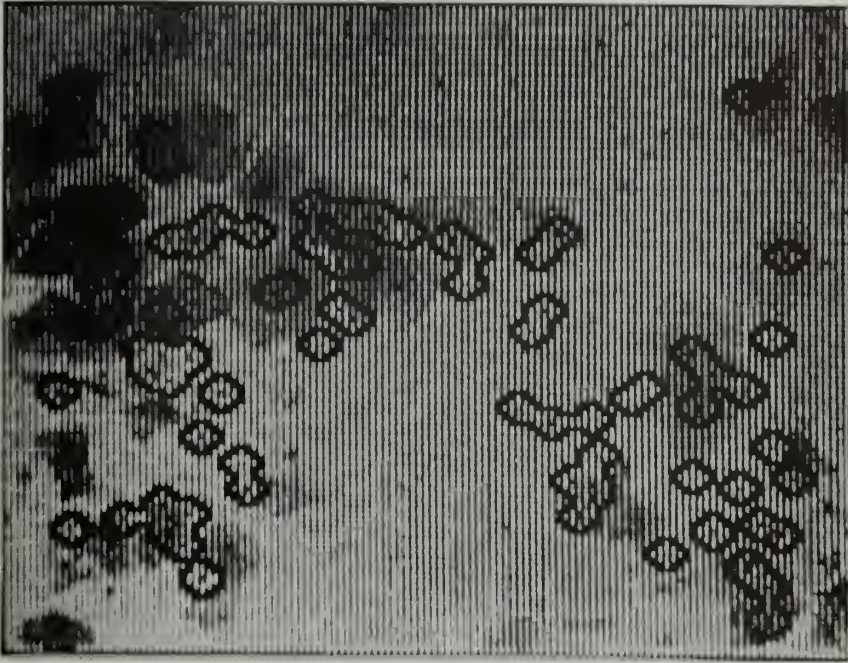


Figure 12(f). Cells Detected

large dark clumps it is said to show an active chromatin texture. This condition can occur in nonmalignant cells, but is much more pronounced in the case of cancer. In Figure 13(b), cell nuclei with active chromatin are approximately centered in the photograph. Note that the drying artifact is a somewhat glassy material containing refractile and light-absorbing areas, while the chromatin textures are mostly composed of light-absorbing areas and are in most cases somewhat less contrasty than the artifact.

To test the applicability of the varivalued logic approach (Section IV.D) to this problem, sample textures were acquired by selecting 14 chromatin and 13 artifact samples from 5 different Pap smears. In some cases the same cell was sampled more than once. An attempt was made to sample textures that appeared to the eye to contain some local texture information and to avoid texture regions that would obviously need contextual data for discrimination. For example, the chromatin in some cell nuclei is so condensed as to present an opaque dark blob. It would make no sense to try to separate these on the basis of local texture.

The 27 samples contained 32×32 picture elements each, with the sample covering an area of perhaps $100 \mu\text{m}^2$ on the microscope slide. The gray values were quantized initially to 16 equally spaced gray levels. This was reduced subsequently to 4 equally spaced gray levels to reduce processing costs.

Nine each of the chromatin and artifact texture samples were designated the training set and the rest were set aside as an unknown or test set. The chromatin texture is considered to be T^1 and the artifact is T^0 .

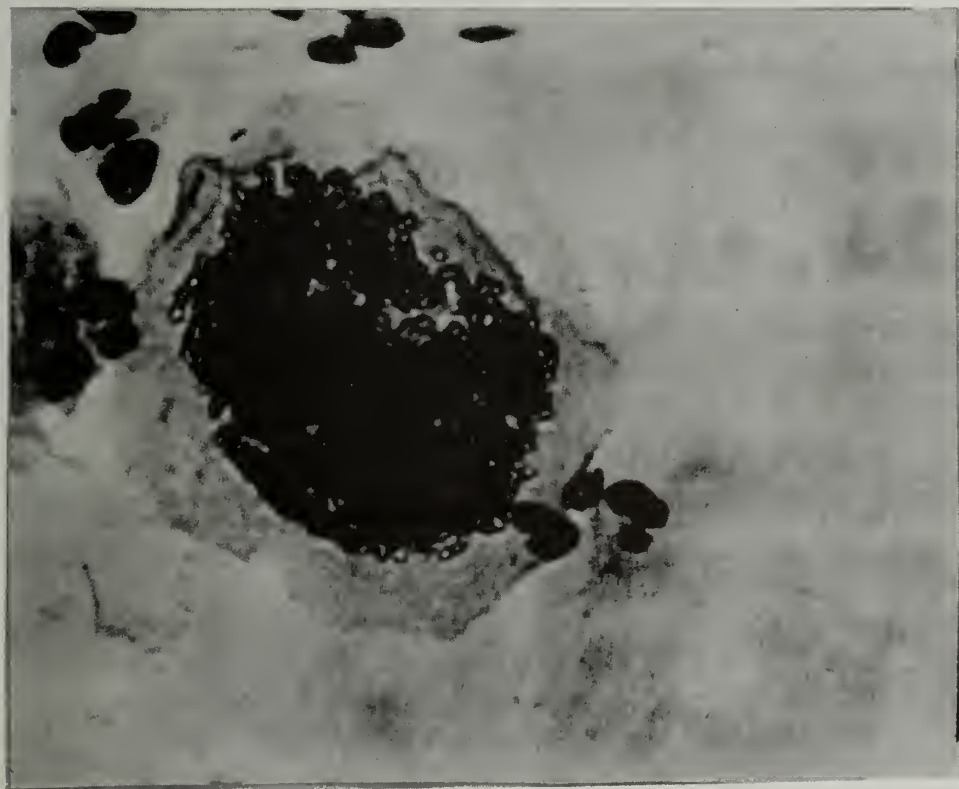


Figure 13(a). Photomicrograph
of Drying Artifact

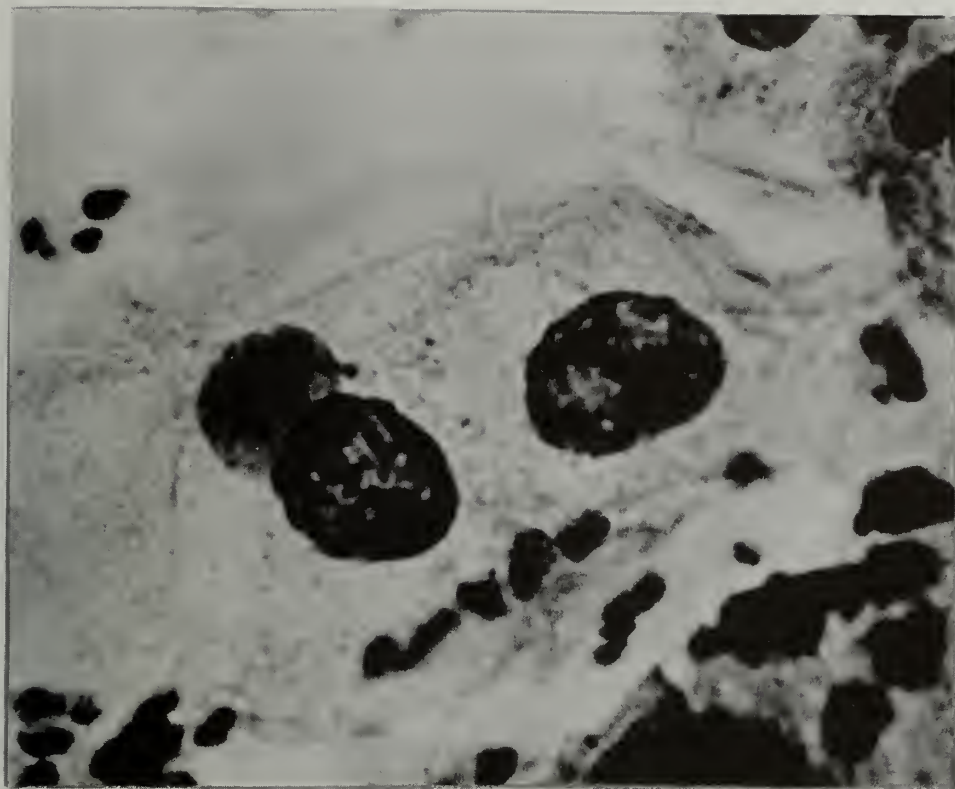


Figure 13(b). Photomicrograph of Cell Nucleus
Showing Chromatin Texture (Center)

A 3×2 neighborhood was used, so there was a total of $30 \times 31 \times 18 = 16,740$ events in the training set. Conditional probabilities and likelihood ratios were computed and the ROC curve labeled D in Figure 13(f) was generated as in the one-dimensional example. According to the ROC, if the decision rule "decide chromatin if $LR(e) > 1$ " is used, one can expect that when a piece of chromatin texture is presented to the categorizer 67 percent of the events will be labeled "chromatin" and when artifact is presented 54 percent of the events will be incorrectly labeled chromatin. This information is then used to set a threshold to classify texture regions, in this case the patches of texture in the samples. (Note that we were previously classifying local neighborhoods. Now we are classifying regions in a digitized picture.) The training set events were labeled using the $LR(e) > 1$ decision rule and the chromatin-labeled events (hits) were counted for each patch of texture. The sample patches were then classified as T^1 (chromatin) if there were more than 650 hits in a patch and T^0 otherwise. There was not complete separation: one-third of the chromatin texture patches were called artifact and 11 percent of the artifact patches were called chromatin, a misclassification rate of about 22 percent on the training set.

Likelihood ratios computed on the training set were used to assign the training set events to T^0 or T^1 , using the $LR > 1$ criterion. These events were fed to the A^Q algorithm and an interval covering which required 31 intervals was generated. When this was applied as a local operation to the unknown texture sample set (Figure 13(c)), the binary result was as shown in Figure 13(d). Black spots in Figure 13(d) indicate that the neighborhood in the vicinity of the spot was categorized as being T^1 by the local operation. These spots were counted for each

Texture patch, as shown in Figure 13(d), and the 650 threshold was used to classify the samples. Figure 13(e) shows how the program classified each sample. In the case of the unknown set, only one error was made.

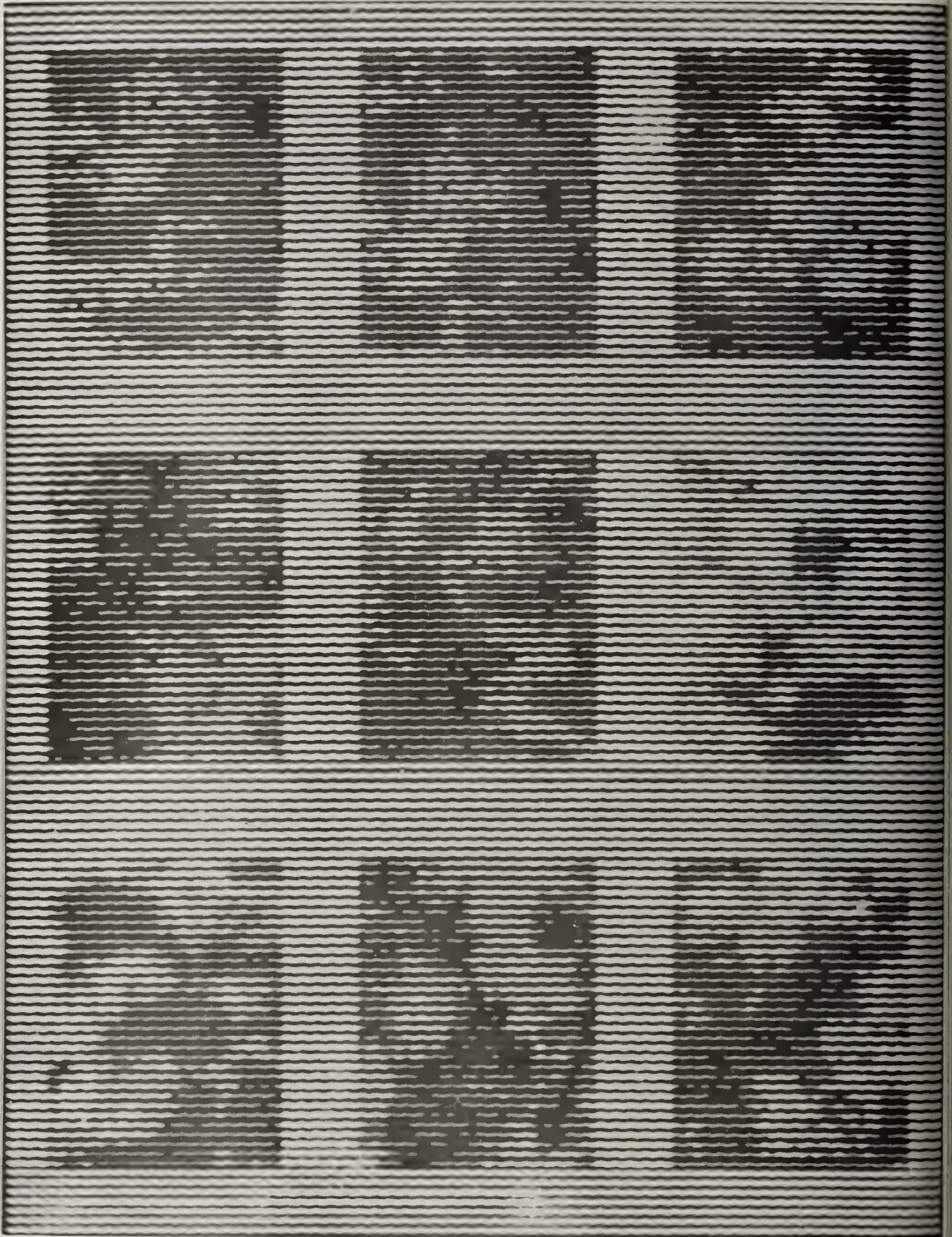


Figure 13(c). Test Set of Texture Samples.

Top Row and First Two in Second Row Are T^1 .

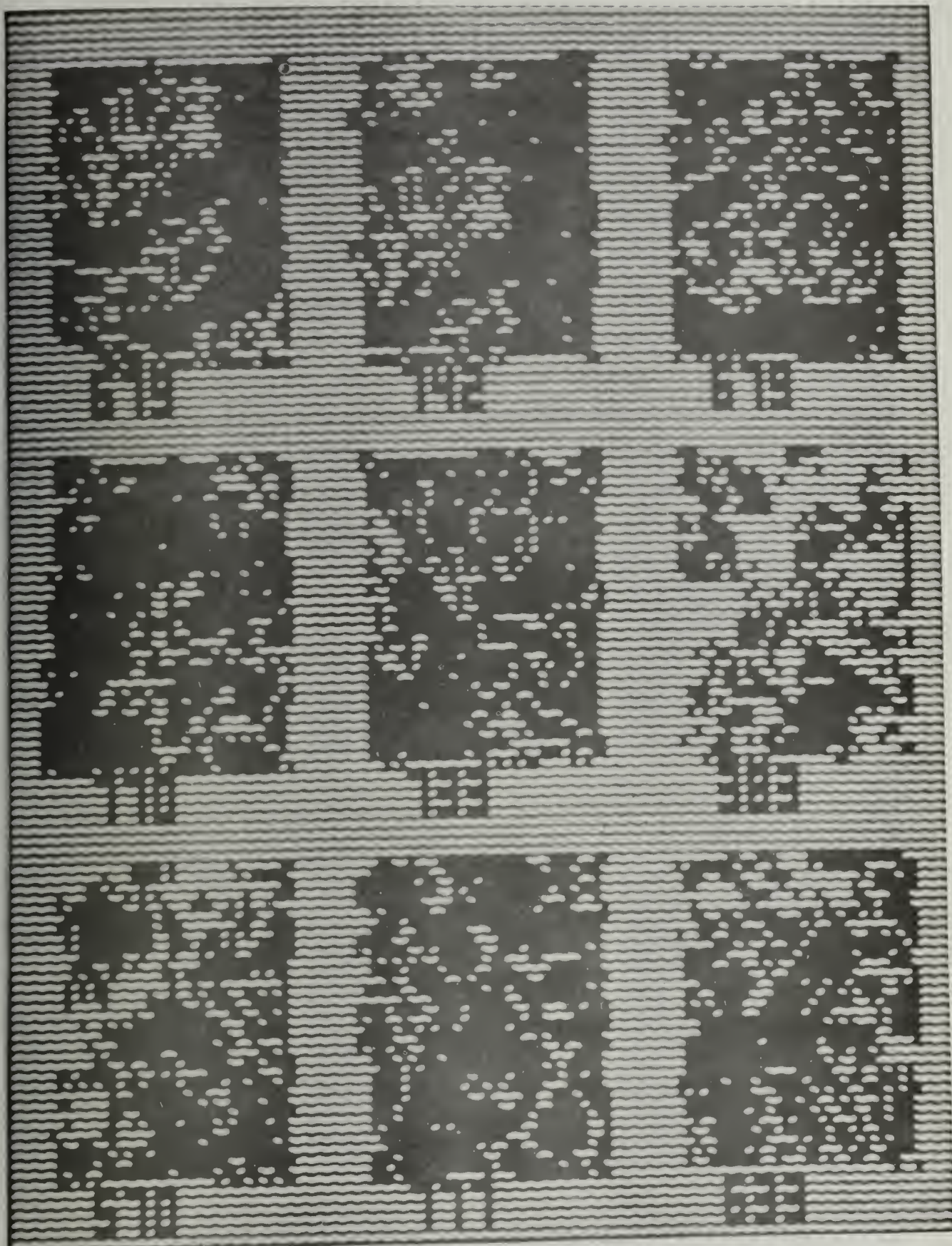


Figure 13(d). Test Set With Hit Counts

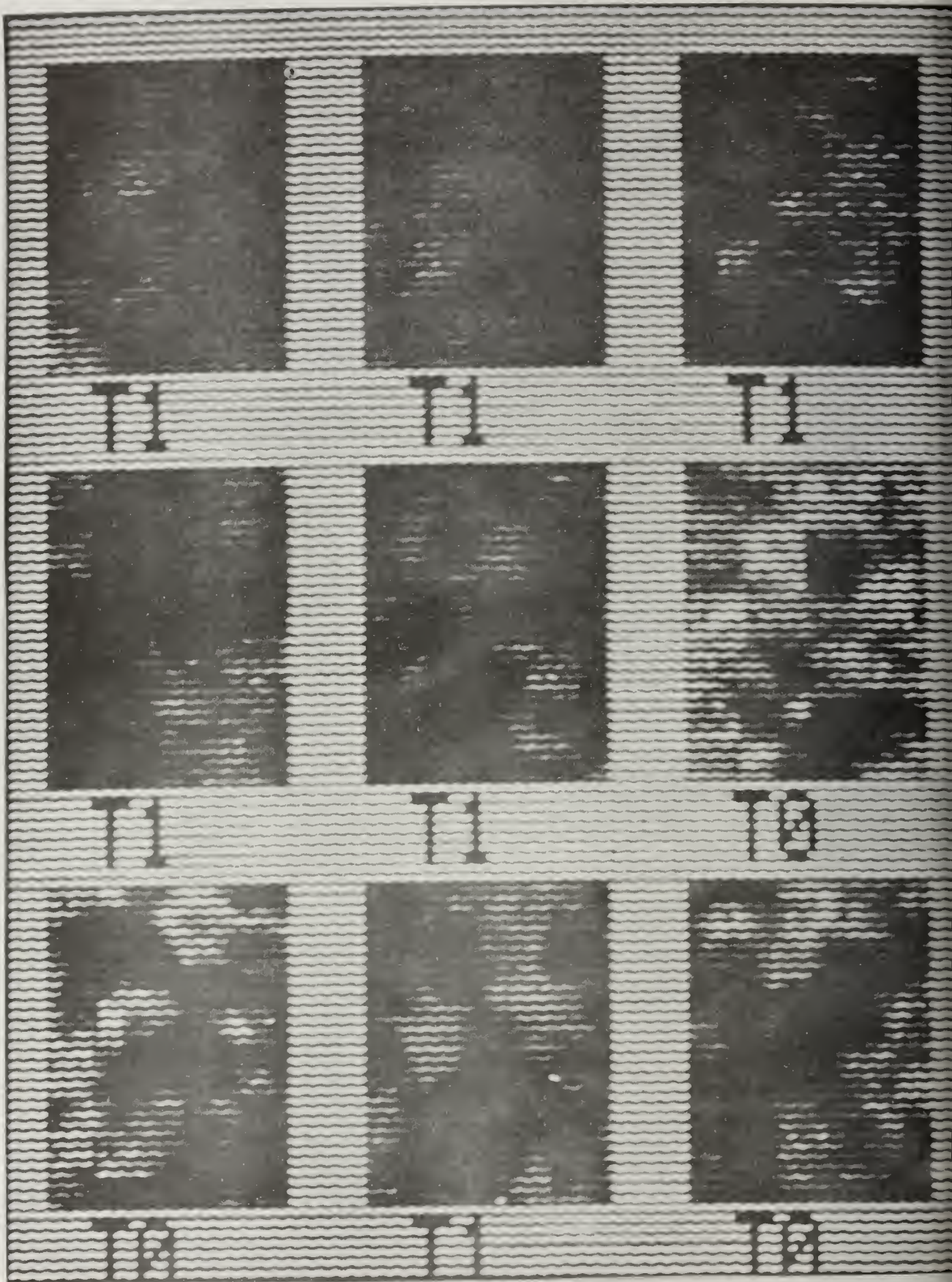


Figure 13(e). Test Set Classified

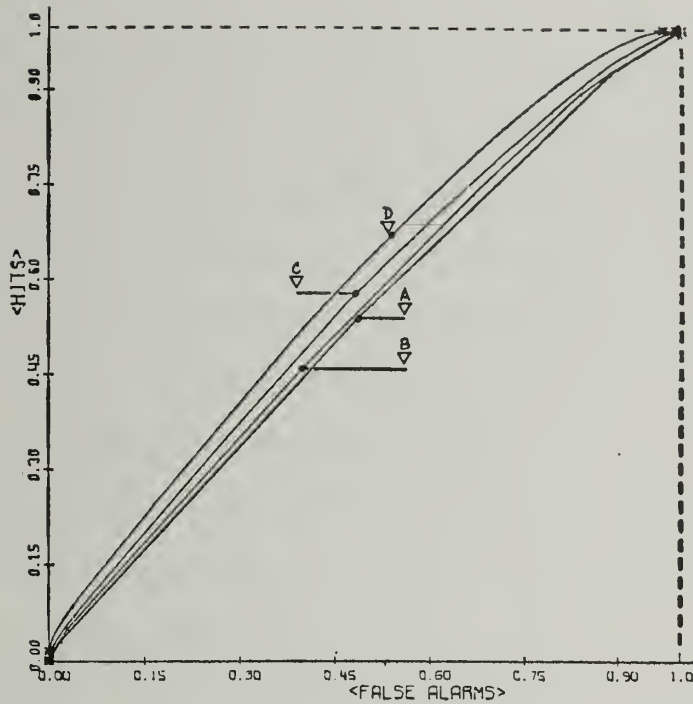


Figure 13(f). Receiver Operating Characteristic for Chromatin vs. Artifact.

Curve A: 1x1 Neighborhood; B: 1x2; C: 2x2; and D: 3x2 Neighborhood.

Triangles indicate performance with $LR(e) \geq 1$ Rule

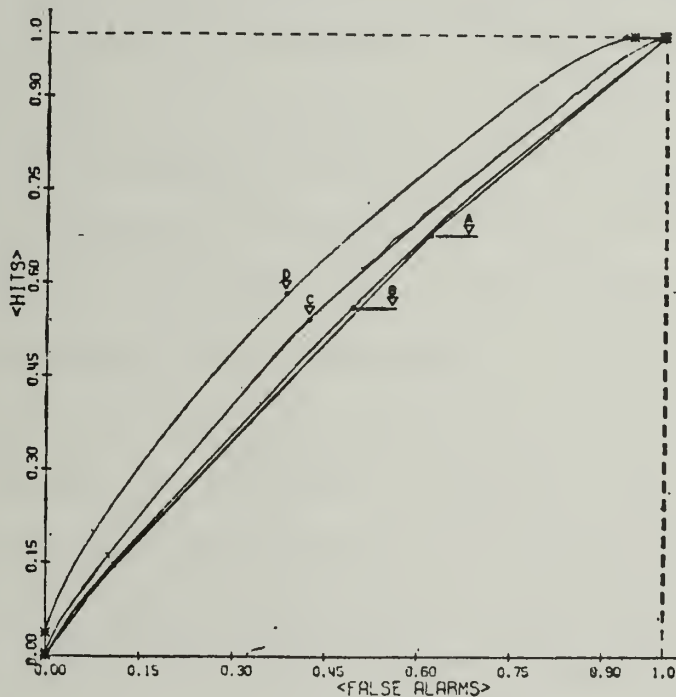


Figure 13(g). Receiver Operating Characteristic with

Tapered Quantization of Samples

VI. CONCLUSIONS

A. Blob and Cell Detection

The performances of the blob and cell detectors indicate that this capability for more complex shape analysis has the potential to overcome the immediate problems which have prevented development of a Cyto-analyzer-like instrument.

The images used in testing the programs were of approximately the level of complexity one would expect in a mechanically dispersed cell sample, such as the one developed for Cytoanalyzer. Therefore, it is reasonable to predict that a processor incorporating the algorithms described above could examine a dispersed cell sample, counting well-differentiated epithelial cells, and detecting enlarged blobs while ignoring white cell clumps, two processes which were essential to the Cytoanalyzer's planned mode of operation, but which proved to be technically infeasible. If the Cytoanalyzer-format slide is used (1 cm x 5 cm), approximately 12,500 fields of 200 μ m x 200 μ m would have to be processed. To do this in four minutes (approximately the time required by human screeners), processing rate would have to approach fifty frames per second. This frame rate is well within the realm of possibility for available television cameras, analog-to-digital converters, and memories at the resolution and signal-to-noise capacities required by the blob-filtering and cell detecting algorithms discussed here. Processing each frame with the algorithms described required 991 fundamental parallel array operations (Boolean operations between planes, unit shifts of planes) and 484 more complex array operations (addition, subtraction and comparison of stacks of planes

interpreted as arrays of binary numbers). To estimate time requirements, it is assumed that each complex operation takes six times as long as a fundamental operation. This rule-of-thumb was a result of observing that integrated-circuit arithmetic/logic units take about six times as long to perform an operation as simple gates take. Using this information and assuming that scanning can be completely overlapped with processing so that all 20 ms are available--it can be calculated that 5.1 μ s are available for each fundamental operation. This execution time is clearly attainable within the present state-of-the-art of circuit technology.

B. Texture Analysis

In view of the success of recent studies in identifying cancer cells through texture features [125, 187] and also in view of the large weight commonly placed on chromatin texture by diagnosticians [52], it seems clear that addition of a texture-analyzing capability would greatly enhance the effectiveness of an image-processing approach to automated cancer cell detection.

The experiments reported here demonstrated that texture information permits a greater degree of discrimination between chromatin and artifact than is obtainable from intensity data alone. However, a degree of ambiguity is not resolved by the present procedure. Figure 13(h) shows the distribution of hit count values (as calculated in Section V.C.) for the sample texture patches in the training set and in the unknown set, a total of 27 observations. While there are too few samples to provide a very reliable picture of the actual distributions, it seems clear that there are two populations and that they overlap. There are several parameters that can be adjusted to try to improve separation: sampling resolution in the initial scan; scanning beam wave length; local neighborhood size and con-

figuration; and quantization scheme. For example, Figure 13(f) shows the result of changing the local neighborhood configuration from 1×1 to 3×2 . The hit probability (for the $LR > 1$ criterion) goes from approximately 55 percent for 50 percent false alarms to 67 percent for 54 percent false alarms, i.e., the hit rate increased 12 percent while the false alarm rate increased 4 percent. The same 16-level texture patches were used in the generation of Figure 13(g); however, an alternative quantization scheme was applied. Instead of quantizing four gray-value ranges of equal width, a "tapered" quantization was used [181]. With this method, the gray-value ranges are adjusted so that approximately the same number of picture elements will have each of the (quantized) values, or to put it another way, each gray value will occur with equal probability in the quantized picture. This has the effect of increasing detail in large areas of low-amplitude high-frequency modulation. The quantization ranges were set once by using a composite gray-value histogram derived from all training set samples. As can be seen by looking at the coordinate of point D in Figure 14(b) for the $LR > 1$ decision rule and using a 3×2 local neighborhood, a 58 percent hit rate can be achieved at the expense of 38 percent false alarms, a separation of 20 percent which is an improvement over the 13 percent separation observed with the equal-range-width quantizing of Figure 13(f). By systematically adjusting parameters in this way and monitoring expected performance via the ROC, it is possible to arrive at an interval cover with optimal effectiveness for this application.

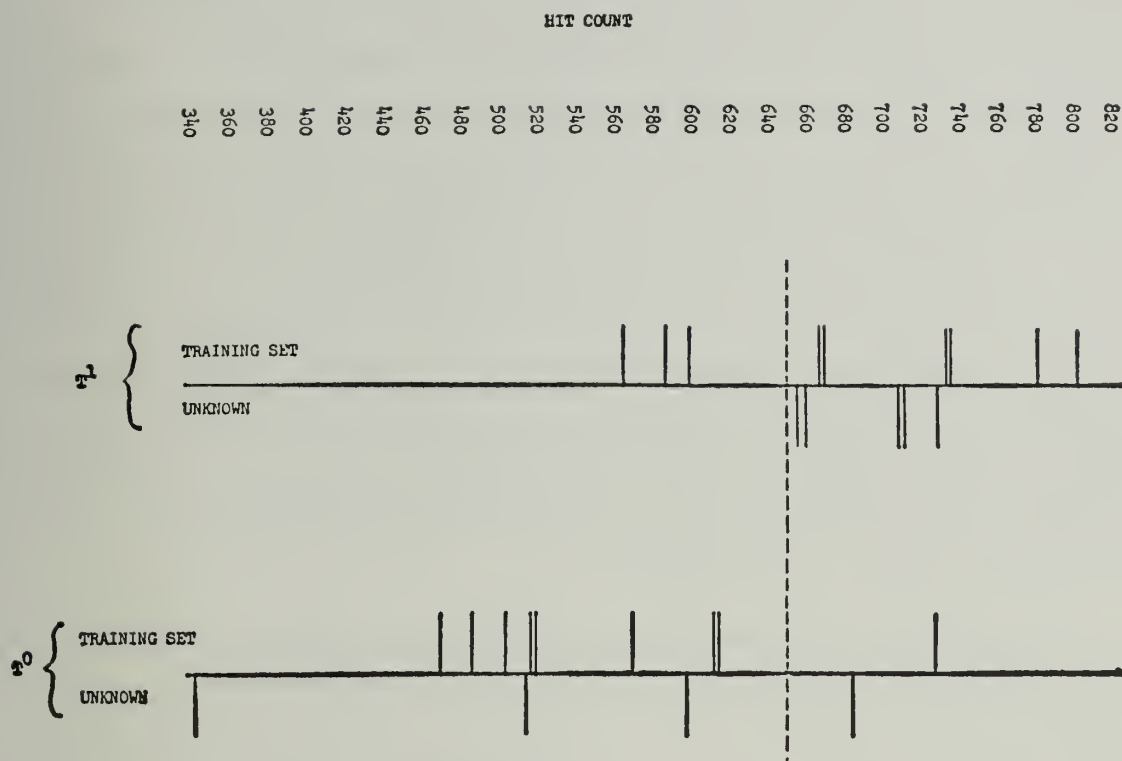


Figure 13(h). Hit Count Distribution

C. Recommendations for Further Work

The results of these experiments, while encouraging, indicate that much work needs to be done in adjusting the system parameters before an acceptable performance level can be achieved. Specifically, studies using the protocols described above should be conducted on large samples of mechanically-dispersed cells so that optimal determinations can be made for variables such as scanning wavelength and photometric and spatial resolution, big-blob detector parameters, and cell detector parameters. To do this economically, the programs must be converted from PAX II to PAU machine language so that existing dedicated image-processing facilities can be used. The IBM 360 is much too expensive to be used for large-scale runs. PAX II is at a slightly higher level than PAU machine language, so the mapping is not one-to-one. However, the basic structure of the algorithm will survive the conversion.

It seems clear that texture, as well as colorimetric data, should be much more extensively investigated as a cell characterizer both for cancer screening and for other cytological image analysis applications. The processing of color data has been well investigated in connection with analysis of multispectral images and data generated by remote sensing devices such as carried by satellites or aircraft. Multispectral data is formatted as a multiple-component vector with each component representing detected energy integrated over some portion of the electromagnetic spectrum. The varivalued logic approach can be used to define efficient color vector categorizers in the same manner as in the texture case. For the cytology application, it is essential that the scanning equipment used be photometrically calibrated so that cytochemical techniques can be employed in devising and evaluating cell analysis programs.

Finally, other applications requiring the extraction of quantitative information from large numbers of cells should be considered. Two particularly interesting applications are in the quantitative study of neuron depopulation and in cell motility investigations. The counting of neurons in sectioned brains is a compelling candidate, since shape analysis is needed to distinguish the neurons from glial cells and because there is a high data rate requirement. The blob-detecting program should be transferable to the neuron counting problem. Motility studies involving location and tracking of moving cells or small organisms in movie film or from real-time television scanning is another application requiring rapid blob location. Tracking individual cells in cell cultures could be of use in following the life history of specific cells being subjected to controlled stresses, e.g., presence of carcinogenic substances.

LIST OF REFERENCES

- [1] Anonymous, "Exfoliative Cytology," reprinted from CA: A Cancer Journal for Clinicians, American Cancer Society.
- [2] Anonymous, "Automated Cytology: A Symposium By Correspondence," *Acta Cytologica*, Vol. 15 (Jan.-Feb.), 1971.
- [3] Anonymous, "Screening for Specific Cytological Features," *World Medicine*, Vol. 3, No. 26, pp. 50-51, 1968.
- [4] Anonymous, "Mass Production Pathology," *World Medicine*, May 20, pp. 52-55, 1969.
- [5] Anonymous, "Automation and Data Processing in Pathology," *Lancet*, Vol. 1 (March 8), pp. 508-509, 1969.
- [6] Anonymous, "Human Blood Cells Counted Accurately by TV Camera Used As 'Eye' of Computer," *Electrical Engineering*, Vol. 73, p. 288, 1954.
- [7] Atkin N. B. and Richards, B. M., "DNA in Human Tumors as Measured by Microspectrophotometry of Feulgen Stain: A Comparison of Tumors Arising at Different Sites," *British Journal of Cancer*, Vol. 10, pp. 769-786, 1956.
- [8] Atkin N. B., "The Deoxyribonucleic Acid Content of Malignant Cells in Cervical Smears," *Acta Cytologica*, Vol. 8, pp. 68-72, 1964.
- [9] Atkin N. B., Mattinson G., and Baker M. C., "A Comparison of the DNA Content and Chromosome Number of Fifty Human Tumors," *British Journal of Cancer*, Vol. 20, pp. 87-101, 1966.
- [10] Atkin N. B., Baker M. C., and Wilson S., "Stemline Karyotypes of 4 Carcinomas of the Cervix Uteri," *American Journal of Obstetrics and Gynecology*, Vol. 99, pp. 506-514, 1967.
- [11] Atkin N.B., "A High Incidence of Cells With a Condensed (Telophase) Chromatin Pattern in Human Tumors and Carcinoma-In-Situ," *Acta Cytologica*, Vol. 11, p. 81, 1967.
- [12] Auersperg N. and Wakonig-Vaartaja T., "Chromosome Changes in Invasive Carcinomas of the Uterine Cervix," *Acta Cytologica*, Vol. 14, pp. 495-501, 1970.
- [13] Bartels, P. H., Wied, G. L., and Bahr G. F., "Cell Recognition from Equiprobable Extinction Range Contours," *Acta Cytologica*, Vol. 12, pp. 205-217, 1968.
- [14] Bartels, P. H., Bahr G. F., and Wied G. L., "Cell Recognition By Cluster Analysis in Pure Parameter Hyperspaces," *Acta Cytologica*, Vol. 12, pp. 371-380, 1968.

- [15] Bartels, P.H., Bahr G.F., Wied G.L., "Cell Recognition From Line Scan Transition Probability Profiles", Acta Cytology, Vol. 13, pp. 210-217, 1969.
- [16] Bartels, P.H., et. al., "Computer Analyses of Lymphocytes in Transformation: A Methodologic Study", Acta Cytology, Vol. 13, pp. 557-568, 1969.
- [17] Bartels P.H., et. al., "A Self-Learning Program for Cell Recognition", Acta Cytology, Vol. 14, pp. 486-494, 1970.
- [18] Bartels P.H., et. al., "Cell Recognition By Neighborhood Grouping Techniques in TICAS", Acta Cytology, Vol. 14, pp. 313-324, 1970.
- [19] Belson M., Dudley, A.W., Ledley, R.S., "Automatic Computer Measurements of Neurons", Pattern Recognition, Vol. 1, pp. 119-128, 1968.
- [20] Bibbo, M., et. al., "Data Bank for Endometrial Cells: Operation of the TICAS File Project", Acta Cytologica, Vol. 14, pp. 574-582, 1970.
- [21] Bieseke, J.J., The Nucleus of the Cancer Cell, Academic Press, New York, 1963.
- [22] Boddington, M.M., Biamond, R.A., "Apparatus for Automatic Preparation and Scanning in Cervical Cytology", British Medical Journal, Vol. 3, pp. 160-161, 1967.
- [23] Borovec, R.T., "The PAX II Picture Processing System at the University of Illinois: Programming Manual", University of Illinois, Department of Computer Science Report No. 314, 1969.
- [24] Bostrom, R.C., Sawyer, H.S., and Tolles, W.E., "Instrumentation for Automatically Pre-Screening Cytological Smears", Proceedings of the IRE, Vol. 47, pp. 1895-1900, 1959.
- [25] Bostrom, R.C., et. al., "The Cytoanalyzer - An Automatic Pre-Screening Instrument for Cancer Detection", Proceedings of the Second International Conference of Medical Electronics, 1959.
- [26] Bostrom, R.C. and Holcomb, W.G., "CYDAC, A Digital Scanning Cytophotometer", IEEE International Convention Record, Vol. 9, pp. 110-119, 1963.
- [27] Brandao, H.J.S., "DNA Content in Epithelial Cells of Dysplasias of the Uterine Cervix. Histologic and Microspectrophotometric Observations", Acta Cytologica, Vol. 13, pp. 232-237, 1969.
- [28] Brecher, G., et. al., "Size Distribution of Erythrocytes", Annals of the New York Academy of Science, Vol. 99, p. 242, 1962.
- [29] Buchold, B.M., etl al., "Erythrocyte Deformability and the Bimodal Volume Distribution", Biophysical Journal, Vol. 8, A-113, 1968.

- [30] Burton, P.A., and Cunliffe, W.J., "A Comparison of Tetracycline Fluorescence and Exfoliative Cytology in the Detection of Malignancy", *Lancet*, pp. 1002-1004, 1966.
- [31] Caspersson, T., Zech, L., and Johansson, C., *Experimental Cell Research*, Vol. 62, p. 490, 1970.
- [32] Clausen, K.P. and Von Haam, E., "Fine Structure of Malignancy-Associated Changes (MAC) in Peripheral Human Leukocytes", *Acta Cytologica*, Vol. 13, pp. 435-442, 1969.
- [33] Cooke-Yarborough and Whyard, R.E., "The Automatic Counting of Red Blood Cells", *British Journal of Applied Physics*, Supplement 3, pp. 147-156, 1954.
- [34] Cornsweet T., Visual Perception, Academic Press, New York, 1970.
- [35] Courtney, W.B., et. al., "Field Trial of the Cytoanalyzer: 1,184 Specimens Analyzed", *Journal of the National Cancer Institute*, Vol. 24, pp. 1167-1186, 1960.
- [36] Courtney, W.B., et. al., "Modified Drop Technique: Preparation of Smears for Automatic, Electronic Scanning", *Journal of the National Cancer Institution*, Vol. 25, pp. 703-711, 1960.
- [37] Crosland-Taylor, P.J., "A Device for Counting Small Particles Suspended in a Fluid Through A Tube", *Nature*, Vol. 171, pp. 37-38, 1953.
- [38] Davidsohn, I., and Louisa, Y.N., "Immunocytology of Cancer", *Acta Cytologica*, Vol. 14, pp. 276-282, 1970.
- [39] Dawson, I.M.P., et. al., "CESAR: Cervical Smear Analyser and Reader", *Journal of Clinical Pathology*, Vol. 20, pp. 724-730, 1967.
- [40] Dell, H.A., "Stages in the Development of an Arrested Scan Type Microscopic Particle Counter", *British Journal of Applied Physics*, Supplement 3, pp. 156-161, 1954.
- [41] Diacumakos, E.G., et. al., "Exfoliated Cell Studies and the Cytoanalyzer", *Annals of the New York Academy of Science*, Vol. 97 pp. 498-513, 1962.
- [42] Diacumakos, P.H., et. al., "Further Developments of the Cytoanalyzer", *Acta Cytologica*, Vol. 6, pp. 238-244, 1962.
- [43] Dupraw, E.J. and Bahr, G.F., "The Arrangement of DNA in Human Chromosomes, as Investigated by Quantitative Electron Microscopy", *Acta Cytologica*, Vol. 13, pp. 188-205, 1969.
- [44] Eaves, "Image Processing in the Biomedical Sciences", *Comput Biomedical Research*, Vol. 1, pp. 112-123, 1967.

- [45] Enslein, K., Data Acquisition and Processing in Biology and Medicine, Vol. 5, Pergamon, New York, 1967.
- [46] Evans, D.M.D. (Ed.), Cytology Automation, Livingstone, London, 1970.
- [47] Evans, D.M.D., Sanerkin, N.G., "Cytology Screening Error Rate", in Cytology Automation, Evnas, D.M.D. (Ed.), Livingstone, London, 1970.
- [48] Fiel, R.J., "Small Angle Light Scattering of Bioparticles. 2. Cells and Organelles", Experimental Cell Research, Vol. 59, pp. 421-428, 1970.
- [49] Fischmeister, H.F., "Automatic Scanning and Measuring Devices", in Stereology: Proceedings of the Second International Congress for Stereology, Elias (Ed.), p. 265, Springer-Verlag, New York, 1967.
- [50] Flory, L.E., "Scanning Microscopy in Biology and Medicine", Proceedings of the IRE, Vol. 47, 1959.
- [51] Friedlander, S., "A New Method for Detecting Changes in the Surface of Human Exfoliated Cervical Cells with the Scanning Electron Microscope", Acta Cytologica, Vol. 13, pp. 288-291, 1970.
- [52] Frost, M.J., The Cell in Health and Disease, Williams and Wilkins, Baltimore, 1969.
- [53] Frost, J.K., "Trichomonas Vaginalis and Cervical Epithelial Changes", Annals of the New York Academy of Sciences, pp. 792-799, 1962.
- [54] Fulwyler, M.J., "Electronic Separation of Biological Cells by Volume", Science, Vol. 156, pp. 910-911, 1965.
- [55] Gallus, G. and Vetter, H.G., "A Family of Computer Processing Procedures for Biological Pictures", Physics in Medicine and Biology, Vol. 15, p. 188, 1970.
- [56] Glaser, D.A. and Wattenburg, W.H., "An Automated System for the Growth and Analysis of Large Numbers of Bacterial Colonies Using An Environmental Chamber and a Computer-Controlled Flying-Spot Scanner", Annals of the New York Academy of Sciences, Vol. 139, pp. 243-257, 1966.
- [57] Glaser, D.A., "An Automated System for Growth and Analysis of Bacterial Colonies", in Image Processing in Biological Science, Ramsey, D. (Ed.), pp. 57-74, 1968.
- [58] Gregg, E.C., Steidley, K.D., "Electrical Counting and Sizing of Mammalian Cells in Suspension", Biophysical Journal, Vol. 5, pp. 393-405, 1965.

- [59] Green, D.M., Swets, J.A., Signal Detection Theory and Psychophysics, Wiley, New York, 1966.
- [60] Green, J.E., "Computer Methods for Erythrocyte Analysis", IEEE Convention Record: Symposium on Feature Extraction and Selection in Pattern Recognition, Argonne, Illinois, October 5-7, pp. 100-109, 1970.
- [61] Grover, B., et. al., "Electrical Sizing of Particles in Suspensions", Biophysical Journal, Vol. 9, pp. 1398-1414, 1969.
- [62] Gusberg, S.B., et. al., "Automated Cytologic Screening for Uterine Cancer: A Feasibility Study", Obstetrics and Gynecology, Vol. 34, 284-287, 1967.
- [63] Ham, A.W., Histology, Lippincott, Philadelphia, 1957.
- [64] Handler, P., Biology and the Future of Man, Oxford University Press, New York, 1970.
- [65] Hawkins, J.K., "Textural Properties for Pattern Recognition", in Picture Processing and Psychopictorics, Lipkin, B.S. and Rosenfeld, A. (Eds.), Academic Press, New York, pp. 347-370, 1970.
- [66] Heanley, C., "Tetronics Trace Laying System", in Cytology Automation, Evans, D.M.D. (Ed.), pp. 154-164, Williams and Wilkins Co., Baltimore, 1968.
- [67] Hennis, R.B., "Recognition of Unnurtured Characters in a Multifont Applications", in Pattern Recognition, Kanal, L.N. (Ed.), pp. 79-100, Thompson Book Co., Washington D.C., 1968.
- [68] Hess, C., "Historical Development of Cytologic Automation", in Summary of State-of-the-Art Workshop on Papanicolaou Smear Analysis, Ramsey-Klee, ed., Department of HEW, NCHS-RD-70-2, 1970.
- [69] Hoffman, J.G., "Flying Spot Methods for Measuring DNA and the Nucleocytoplasmic Volume Ratio in Tissue Cells", Annals of the New York Academy of Science, Vol. 97, pp. 380-394, 1962.
- [70] Holland, P. and Sleamaker, K., "A Cytologic Evaluation of Intracellular Tetracycline Fluorescence in Normal and Neoplastic Lymphocytes", Acta Cytologica, Vol. 13, pp. 246-249, 1969.
- [71] Horvath, W.J., et. al., "Quantitative Measurements of Cell Properties in Papanicolaou Smears as Criteria for Screening", Transactions of the First International Cancer Cytological Congress, pp. 371-397, 1956.
- [72] Hrushovetz, S.B., "Two-Wavelength Feulgen Cytophotometry of Cells Exfoliated From the Uterine Cervix", Acta Cytologica, Vol. 13, pp. 583-594, 1969.

- [73] Hulett, H.R., et. al., "Cell Sorting: Automated Separation of Mammalian Cells As a Function of Intracellular Fluorescence", *Science*, Vol. 166, pp. 747-749, 1969.
- [74] Ingram, M. and Preston, K., "Importance of Automatic Pattern Recognition Techniques in the Early Detection of Altered Blood-Cell Production", *Annals of the New York Academy of Science*, Vol. 113, pp. 1066-1072, 1964.
- [75] Ingram, M., Norgren, P.E., Preston, K., "Automatic Differentiation of White Blood Cells", in Image Processing in Biological Science, Ramsey, D. (Ed.), pp. 97-117, 1968.
- [76] Ishiyama, T., et. al., "A Study of the Automation of Cytodiagnosis", *Medical and Biological Engineering*, Vol. 7, pp. 297-306, 1969.
- [77] Izzo, N.F. and Coles, W., "Blood Cell Scanner Identifies Rare Cells", *Electronics*, Vol. 35, pp. 52-57, 1962.
- [78] James, T.W. and Anderson, N.G., "Continuous Recording of Cell Number in Logarithmic and Synchronized Cultures", *Science*, Vol. 142, pp. 1183-1185, 1963.
- [79] Johnston, B. and Brady, J.M., "Malignancy Related Changes in Peripheral Blood Smears", *Acta Cytologica*, Vol. 14, pp. 399-403, 1970.
- [80] Johnston, E.G., "The PAX II Picture Processing System", in Picture Processing and Psychopictorics, Lipkin, B.S. and Rosenfeld, A. (Eds.), Academic Press, New York, pp. 427-512, 1971.
- [81] Johnstone, M.C. (ed.), "The Cervix: Proceedings of a Conference of the New York Academy of Science", *Annals of the New York Academy of Science*, Vol. 97, pp. 527-878, 1962.
- [82] Kamentsky, L.A., "Pattern and Character Recognition Systems - Picture Processing by Nets of Neuron-Like Elements", in Proceedings of the Western Joint Computer Conference, San Francisco, p. 304, 1959.
- [83] Kamentsky, L.A., Berman, H., and Melamed, M., "Ultraviolet Absorption in Epidermoid Cancer Cells", *Science*, Vol. 142, p. 1580, 1963.
- [84] Kamentsky, L.A., "Rapid Biological Cell Identification By Spectroscopic Analysis", in Proceedings of the 18th Annual Conference on Engineering in Medicine and Biology, Vol. 7, Frommer, P.L., and Vurek, G.G. (Eds.), p. 178, 1965.
- [85] Kamentsky, L.A., Melamed, M.R., and Berman, H., "Spectrophotometer: New Instrument for Ultrarapid Cell Analysis", *Science*, Vol. 150, pp. 630-631, 1965.

- [86] Kamentsky, L.A. and Melamed, M.R., in the Annals of the New York Academy of Science, 1967.
- [87] Kamentsky, L.A., "Cell Identification and Sorting", in Computers in Biomedical Research, Vol. 3, Stacy, R.W. and Waxman, B.D. (Eds.), pp. 107-144, 1969.
- [88] Kamentsky, L.A., Melamed, M.R., "Rapid Multiple Cellular Mass Constituent Analysis", Annals of the New York Academy of Science, Vol. 157, pp. 310-323, 1969.
- [89] Kamentsky, L.A. and Thorell, B., "Cell Population Identification Studies", Acta Cytologica, Vol. 14, pp. 307-312, 1970.
- [90] Kamentsky, L.A., "Improvements in a Rapid Cell Spectrophotometer for Higher Resolution Cell Identification", in Proceedings of the Third International Congress of Histochemistry and Cytochemistry, New York, Springer, Berlin, 1968.
- [91] Kamentsky, L.A., and Melamed, M.R., "Spectrophotometric Cell Sorter", Science, Vol. 156, pp. 1364-1365, 1967.
- [92] Kawamura, A. (Ed.), Fluorescent Antibody Techniques and Their Application, University of Tokyo Press, Tokyo, 1969.
- [93] Kirsch, R., "Computer Determination of Constituent Structure of Biological Images", National Bureau of Standards Report 10, Vol. 173.
- [94] Koenig, S.H., Brown, R.D., et. al., "Efficacy of a Rapid Cell Spectrophotometer in Screening for Cervical Cancer", Cancer, Vol. 21, pp. 1019-1026, 1968.
- [95] Kopac, M.J. and Harris, J., "Microsurgery and Visible Light Television", Annals of the New York Academy of Science, Vol. 97, pp. 331-345, 1962.
- [96] Kubitschek, H.E., "Electronic Counting and Sizing of Bacteria", Nature, Vol. 182, pp. 234-235, 1958.
- [97] Ladinsky, J.L., Sario G.E., and Peckham, B.M., "Cell Size Distribution Patterns as a Means of Uterine Cancer Detection", Journal of Laboratory Clinical Medicine, Vol. 64, p. 970, 1964.
- [98] Lagercrantz, C., "Photoelectric Counting of Individual Microscopic Plant and Animal Cells", Nature, Vol. 161, p. 25, 1948.
- [99] Lagerlof, B. and Adams, L.R., "Cytophotometric Study of Cervical Smears", Acta Cytologica, Vol. 13, pp. 57-58, 1969.
- [100] Lappin, P.R., "Evaluation of Fisher Autocytometer", Laboratory Practices, Vol. 19, pp. 907-908.

- [101] Ledley, R.S., "High-Speed Automatic Analysis of Biomedical Pictures", Science, Vol. 146, pp. 216-223, 1964.
- [102] Ledley, R.S., and Ruddle, F.H., "Chromosome Analysis By Computer", Scientific American, Vol. 214, pp. 40-46, 1966.
- [103] Ledley, R.S., "Automatic Pattern Recognition for Clinical Medicine," Proceedings of the IEEE, Vol. 57, 1969.
- [104] Leif, R.C., "Buoyant Density Separation of Cells", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 21-96, 1970.
- [105] Leif, R.C., "A Proposal for an Automatic Multiparameter Analyzer for Cells (AMAC)", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 131-159, 1970.
- [106] Levialdi, S., "Parallel Pattern Processing", IEEE Transactions on Systems, Man and Cybernetics, October, pp. 394-396, 1971.
- [107] Lipkin, L.E., Watt, W.C. and Kirsch, R.A., "The Analysis, Synthesis, and Description of Biological Images", Annals of the New York Academy of Science, Vol. 128, pp. 984-1012.
- [108] Lubs, H.A. and Ruddle, F.H., "Applications of Quantitative Karyotype to Chromosome Variation in 4400 Consecutive Newborns", in Human Population Cytogenetics, Jacobs, P. (Ed.), 1970.
- [109] Lubs, H.A., et. al., "Display of Chromosome Statistics (DDCS)", in Computers in Biology and Medicine, 1970.
- [110] Malmgren, R.A., "A Cytopathologists Criteria for Malignant Cell Recognition", in Summary of State-of-the-Art Workshop on Papanuolaon Smear Analysis, Ramsey, K.D. (Ed.), Department of Health, Education and Welfare, 1970.
- [111] Mansberg, H.P., et. al., "Spot Scanner Counts Micron-Sized Particles", Electronics, Vol. 30, pp. 142-146, 1957.
- [112] Mansberg, H.P., "Automatic Particle and Bacterial Colony Counter", Science, Vol. 126, pp. 823-827, 1957.
- [113] Mansberg, H.P., "Photo-Optical Field Measurement By Flying-Spot Scanners for Biomedical Applications", SPIE Journal, Vol. 2, pp. 83-88, 1964.
- [114] Mansberg, H.P., "Improved Optical Scanning Devices for Data Acquisition From Biomedical Images", in Data Acquisition and Processing in Biology and Medicine, Vol. 5, Enslein, K. (Ed.), pp. 173-194, 1968.
- [115] Mansberg, H.P. and Segarra, J.M., "Counting of Neurons By Flying Spot Microscope", Annals of the New York Academy of Science, Vol. 99, pp. 309-322, 1962.

- [116] Mansberg, H.P. and Kusnetz, J., "Quantitative Fluorescence Microscopy: Fluorescent Antibody Scanning Techniques", *Journal of Histochemistry and Cytochemistry*, Vol. 14, No. 3, pp. 260-273, 1966.
- [117] Masin, F. and Masin M., "Cell Distribution Technique", *Acta Cytologica*, Vol. 13, pp. 541-542, 1969.
- [118] Mattern, C.F.T., et. al., "Determination of Number and Size of Particles By Electrical Gating: Blood Cells", *Journal of Applied Physiology*, Vol. 10, pp. 56-70, 1957.
- [119] Mattson, J.C. and Von Haam, E., "The Presence of 'Malignancy Associated Changes' in the Monocytes of the Peripheral Blood of Cancer Patients", *Acta Cytologica*, Vol. 11, pp. 308-311, 1967.
- [120] Mawdesley-Thomas, L.E. and Healey, P., "Automated Analysis of Cellular Change in Histological Sections", *Science*, Vol. 163, p. 1200, 1969.
- [121] Mayall, B.H., "Deoxyribonucleic Acid Cytophotometry of Stained Human Leukocytes - I: Differences Among Cell Types", *Journal of Histochemistry and Cytochemistry*, Vol. 17, pp. 249-257, 1968.
- [122] McCarthy, B.D. and Rosen D., "A Machine for the Automatic Extraction of Cell Parameters From Microscope Slide Preparations", *Physics in Medicine and Biology*, Vol. 15, p. 189, 1970.
- [123] McCormick, B.H., "The Illinois Pattern Recognition Computer - Illiac III, *IEEE Transactions on Electronic Computers*, Vol. E-12, No. 5, pp. 791-813, 1963.
- [124] McMaster, G.W., "Quantitative Scanning of Cervical Cells", *Journal of Obstetrics and Gynecology of the British Empire*, Vol. 72, pp. 936-939, 1965.
- [125] McMaster, G.W., "A Measurement of the Pattern of Normal and Malignant Cervical Cells", *Acta Cytologica*, Vol. 12, pp. 9-14, 1968.
- [126] Melamed, M.R., Kametsky, L.A., and Boyse, E.A., "Cytotoxic Test Automation: A Live-Dead Cell Differential Counter", *Science*, Vol. 163, pp. 285-286, 1969.
- [127] Mellors, R.C. and Silver, R., "A Microfluorometric Scanner for the Differential Detection of Cells: Application to Exfoliative Cytology", *Science*, Vol. 114, pp. 356-360, 1951.
- [128] Mellors, R.C., Keane, J.F., and Papanicolaou, G.N., in *Science*, Vol. 116, p. 265, 1952.
- [129] Mellors, R.C., Glassman, A., and Papanicolaou, G.N., "A Microfluorometric Scanning Method for the Detection of Cancer Cells in Smears of Exfoliated Cells", *Cancer*, Vol. 5, pp. 458-468, 1952.

- [130] Mellors, R.C., "Fundamental Studies of the Cell: Observations Underlying the Concept of Cancer-Cell Detection by Automatic Scanning Microscopes", in The Laboratory Diagnosis of Cancer of the Cervix, Homburger and Fishman (eds.), pp. 23-32, 1956.
- [131] Mendel, J.M., Fu K.S., Adaptive Learning and Pattern Recognition Systems, Academic Press, New York, 1970.
- [132] Mendelsohn, M.L., Mayall, B.H., and Prewitt, J.M.S., "Approaches to the Automation of Chromosome Analysis", in Image Processing in Biological Science, Ramsey, D. (ed.), University of California Press, Berkeley, 1968.
- [133] Mendelsohn, M.L., Kolman, W.A., and Bostrom R.C., "Initial Approaches to the Computer Analysis of Cytophotometric Fields", Annals of the New York Academy of Science, Vol. 115, pp. 993-1009, 1964.
- [134] Mendelsohn, M.L., et. al., "Digital Transformations and Computer Analysis of Microscopic Images", in Advances in Optical and Electron Microscopy, Coslett, V.E. (Ed.), pp. 70-150, 1968.
- [135] Meyer, E.W., "Vickärs' Continuous Film", in Cytology Automation, Evans, D.M.D. (Ed.), pp. 147-153, Williams and Wilkins Co., Baltimore, MD, 1968.
- [136] Michalski, R.S., "A Geometrical Model for the Synthesis of Interval Covers", University of Illinois, Department of Computer Science Report No. 461, 1971.
- [137] Michalski, R.S. and McCormick, B.H., "Interval Generalizations of Switching Theory", University of Illinois, Department of Computer Science Report No. 442, 1971.
- [138] Minsky, M., "Steps Toward Artificial Intelligence", in Computers and Thought, Feigenbaum, E.A. and Feldman, Je. (Ed.), pp. 406-450, McGraw-Hill, New York, 1963.
- [139] Minsky, M., "Artificial Intelligence and Intelligent Automata", Project MAC Progress Report VI, pp. 11-54, 1969.
- [140] Minsky, M., Papert, S., Perceptrons, MIT Press, Cambridge, 1969.
- [141] Montgomery, P.O., "Simultaneous Ultraviolet and Visible Light Flying Spot Television Microscopy", Annals of the New York Academy of Science, Vol. 97, p. 490, 1962.
- [142] Moore, G.A., "Application of Computers to Quantitative Analysis of Microstructure", in Stereology: Proceedings of the Second International Congress for Stereology, Elias (Ed.), Springer-Verlag, New York, 1967.

- [143] Mullaney, P.F. and Dean, P.N., "Cell Sizing: A Small-Angle Light-Scattering Method for Sizing Particles of Low Relative Refractive Index", *Applied Optics*, Vol. 8, pp. 2361-2362, 1969.
- [144] Mullaney, P.F., et. al., "Cell Sizing: A Light Scattering Photometer for Rapid Volume Determination", *Review of Scientific Instruments*, Vol. 40, pp. 1029-1032, 1969.
- [145] Mullany, P.F., "The Small Angle Light Scattering of Biological Cells: Theoretical Considerations", *Biophysics Journal*, Vol. 10, pp. 764-772, 1970.
- [146] Mumma, V.R., et. al., "A Particle Size analyzer for Aerosols", *Annals of the New York Academy of Science*, Vol. 99, pp. 298-308, 1962.
- [147] Nadel, E.M., "Computer Analysis of Cytometric Fields By CYDAC and its Historical Evolution From the Cytoanalyzer", *Acta Cytologica*, Vol. 9, pp. 203-205, 1965.
- [148] Nagy, G., "State of the Art in Pattern Recognition", *Proceedings of the IEEE*, Vol. 56, pp. 836-862, 1968.
- [149] Narasimhan, R., "A Programming Language for the Parallel Processing of Pictures", *University of Illinois, Digital Computer Laboratory Report No. 152*, January 1963.
- [150] Narasimhan, R., "Syntactic Descriptions of Pictures and Gestalt Phenomena of Visual Perception", *University of Illinois, Urbana Illinois, Digital Computer Laboratory Report No. 142*, July 1963.
- [151] Narasimhan, R., "A Linguistic Approach to Pattern Recognition", *Digital Computer Laboratory, University of Illinois, Urbana, Illinois Report No. 121*, July 1962.
- [152] Narasimhan, R., "Labeling Schemata and Syntactic Descriptions of Pictures", *Information and Control*, Vol. 7, pp. 151-179, 1964.
- [153] Nasiell, M., "Histiocytes and the Histiocytic Reaction in Vaginal Cytology: A Survey of the Risks of False Diagnosis", *Cancer*, Vol. 14, No. 6, pp. 1223-1225, 1961.
- [154] Nieburgs, H.E., "Recent Progress in the Interpretation of Malignancy Associated Changes (MAC)", *Acta Cytologica*, Vol. 12, pp. 445-453, 1968.
- [155] O'Brien, R. and Rosendranz, P.G., "Studies of Stained Cell Populations with the Rapid Cell Spectrophotometer", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 123-130, Academic Press, New York, 1970.
- [156] Papanicolaou, G.N. and Traut, H., "Diagnosis of Uterine Cancer by the Vaginal Smear", *The Commonwealth Fund*, New York, 1943.

- [157] Papanicolaou, G.N., "A Survey of the Actualities and Potentialities of Exfoliative Cytology in Cancer Diagnosis", *Annals of Internal Medicine*, Vol. 31, pp. 661-674, 1949.
- [158] Papanicolaou, G.M., Atlas of Exfoliative Cytology, Harvard University Press, Cambridge, Massachusetts, 1954.
- [159] Paton, K., "An Automatic Method for Finding Metaphase Spreads", in Pictorial Pattern Recognition, Cheng, G.C., Ledley, R.S., et. al., (Eds.), pp. 135-136, 1966.
- [160] Postlethwaite and Mills, "An Automated Device for Cinematography of Sequential Microscope Sections", in Stereology: Proceedings of the Second International Congress for Stereology, Elias (Ed.), Springer-Verlag, New York, 1967.
- [161] Preston, K., "The Cellscan System - A Leukocyte Pattern Analyzer", *Proceedings of the WJCC*, Vol. 173, 1961.
- [162] Preston, K., "Machine Techniques for Automatic Leukocyte Pattern Analysis", *Annals of the New York Academy of Science*, Vol. 97, pp. 482-490, 1962.
- [163] Preston, K., "Automation of the White Blood Cell Differential Count", *Scientific American*, November, pp. 72-82, 1970.
- [164] Preston, K., "Feature Extraction by Golay Hexagonal Pattern Transforms", *IEEE Transactions on Computers*, Vol. C-20, pp. 1007-1013, 1971.
- [165] Preston, K., "Transcript of Workshop Report on Automatic Cytology", Engineering Foundation Research Conference on Automatic Multiphasic Health Testing, Davos, Switzerland, September 14-18, 1970.
- [166] Prewitt, J.M.S. and Mendelsohn, M.L., "The Analysis of Cell Images", *Annals of the New York Academy of Science*, Vol. 128, pp. 1035-1053, 1966.
- [167] Prewitt, J.M.S. and Mendelsohn, M.L., "The Analysis of Cell Images, II.", *Annals of the New York Academy of Science*, 1968.
- [168] Prewitt, J.M.S., "The Selection of Sampling Rate for Digital Scanning", *IEEE Transactions BME*, Vol. 12, pp. 14-21, 1965.
- [169] Ramsey-Klee, D. (Ed.), "State-of-the-Art Workshop on Papanicolaou Smear Analysis", Department of Health, Education and Welfare, NCHS-RD-70-2, 1970.
- [170] Ramsey-Klee, D. (Ed.), Image Processing in Biological Science, University of California Press, Berkeley, 1968.
- [171] Read, J.S., "Show-and-Tell: An Interactive Programming System for Image Processing", Department of Computer Science, University of Illinois, Urbana, Illinois Report No. 429, 1971.

- [172] Read, J.S. and Jayaramamurthy, S.M., "Automatic Generation of Texture Feature Detectors", IEEE Transactions on Computers, July, 1972.
- [173] Reagan, J.W., Patten, S.F., Atlas of Cytology: The Female Reproductive Tract, American Society of Clinical Pathologists, Chicago, 1966.
- [174] Reid, B.L. and Singh, S., "DNA Values (Feulgen Microspectrophotometry) in Epithelium of Human Ectocervix, Normal and Cancerous", Journal of the National Cancer Institute, Vol. 25, pp. 1291-1301, 1960.
- [175] Rimmer, D., "The Feasibility of Automating Pre-Screening of Cervical Smears on the Basis of Nuclear Size", Journal of Obstetrics and Gynaecology of the British Commonwealth, Vol. 75, pp. 1045-1048, 1968.
- [176] Rimmer, D., "Report on the Performance of the Vickers Cytology Screening Apparatus, Journal of Obstetrics and Cynaecology of the British Commonsealth, Vol. 76, pp. 1029-1034, 1969.
- [177] Rintala, W.M. and Hsu, C.C., "A Feature-Detection Program for Patterns With Overlapping Cells", IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, pp. 16-23, 1968.
- [178] Roberts, F. and Young, J.Z., "The Flying-Spot Microscope", Proceedings of the British Institute of Electrical Engineering, Vol. 741, 1952.
- [179] Rosenberg, S.A., Ledeen, K.S., and Kline, T., "Automatic Identification and Measurement of Cells by Computer", Science, Vol. 163, pp. 1065-1066, 1969.
- [180] Rosenfeld, A., "Noise Cleaning in Binary-Valued Digital Pictures Using 'Propagation' Processes", University of Maryland Computer Science Center, Technical Report 69-91, 1969.
- [181] Rosenfeld, A., Picture Processing by Computer, Academic Press, New York, 1969.
- [182] Rosenfeld, A., and Thurston, M., "Edge and Curve Detection for Visual Scene Analysis," IEEE Transactions on Computers, Vol. C-20, No. 5, pp. 562-569, 1971.
- [183] Ruch, F., "Principles and Some Applications of Cytofluormetry", in Introduction to Quantitative Cytochemistry - II, Wied, G.L. and Bahr, G.F. (Eds.), p. 431, 1970.
- [184] Russel, J.R., Stevens, M., and Boone, M.L.M., "In Vivo Microscopic Holography", Physics in Medicine and Biology, Vol. 15, p. 191, 1970.

- [185] Sandritter, W., Lobel, B.L., and Kiefer, G., "Photometric Cyto-diagnosis of Vaginal Smears", Journal of the National Cancer Institute, Vol. 32, pp. 1221-1228, 1964.
- [186] Sandritter, W., Carl, M., and Ritter, W., "Cytophotometric Measurements of the DNA Content of Human Malignant Tumors by Means of the Feulgen Reaction", Acta Cytologica, Vol. 10, pp. 26-30, 1966.
- [187] Sandritter, W. and Kiefer, G., "Objectivization of Chromatin Patterns Using the Fast-Scanning Stage of the UMSP-1", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G. (Ed.), pp. 177-185, 1970.
- [188] Sawyer, H.S. and Bostrom, R.C., "A New Nipkow Disk-Scanner for Accurate Cytological Measurement", 1958 IRE Convention, Medical Electronics, IRE National Convention Record, Vol. 9, p. 37, 1958.
- [189] Sebestyan, G.S., Decision-Making Processes in Pattern Recognition, MacMillan, New York, 1962.
- [190] Sebestyan, G.S., "Machine Aided Reconnaissance Photointerpretations", SPIE Journal, Vol. 2, pp. 89-92, 1963.
- [191] Spencer, C.C. and Bostrom R.C., "Performance of the Cytoanalyzer in Recent Clinical Trials", Journal of the National Cancer Institute, Vol. 29, No. 2, 1962.
- [192] Spriggs, A.I., "Automated Screening for Cervical Smears", Lancet, Vol. I, pp. 359-360, 1968.
- [193] Stein, P.G., Lipkin, L.E., and Shapiro, H.M., "Spectre II: General-Purpose Microscope Input for a Computer", Science, Vol. 166, p. 3930.
- [194] Steward, V.W., "The Use of a New Type of Computerized Flying-Spot Microscope for the Evaluation of Pathological Material", Physics in Medicine and Biology, Vol. 15, p. 189, 1970.
- [195] Sutton; R.N. and Hall, E.L., "Texture Measures for Automatic Machine Recognition and Classification of Pulmonary Disease", Proceedings, Two Dimensional Signal Processing Conference, October 6-8, 1971, Columbia, Missouri, 1971.
- [196] Taylor, W.S., "A Versatile Cell Detector for Cell Volume Measurements", Medical Biological Engineering, Vol. 8, pp. 281-290, 1970.
- [197] Taylor, W.K., "An Automatic System for Obtaining Particle Size Distribution with the Aid of a Flying Spot Microscope", British Journal of Applied Physics, Supplement 3, pp. 173-175, 1954.
- [198] Tolles, W.E., "A Multi-Dimensional Analysis of Some Quantitative Characteristics of Exfoliated Cells in Papanicolaou Smears", Transactions of the Annual Inter-Society Cytology Council, 1954.

- [199] Tolles, W.E., "The Cytoanalyzer - An Example of Physics in Medical Research", Transactions of the New York Academy of Science, Vol. 17, pp. 250-256, 1955.
- [200] Tolles, W.E. and Bostrom, R.C., "Automatic Screening of Cytological Smears - the Instrumentation", Annals of the New York Academy of Science, Vol. 63, pp. 1211-1218, 1956.
- [201] Tolles, W.E., "The Automatic Scanner", Transactions of the First International Cancer Congress, pp. 211-219, 1956.
- [202] Tolles, W.E., et. al., "The Application of Automatic High Speed Measurement Techniques to Cytology", IRE Convention Record, Part 9, pp. 17-23, 1956.
- [203] Tolles, W.E. and Bostrom, R.C., "Automatic Scanning of Cytological Smears for Cancer", Annals of the New York Academy of Science, Vol. 63, p. 1211, 1956.
- [204] Tolles, W.E. and Horvath, W.J., "An Examination of the Value of Some Quantitative Characteristics of Exfoliated Cells as a Means of Screening for Female Genital Cancer", Transactions of the Inter-Society Cytology Council, pp. 93-99, 1957.
- [205] Tolles, W.E., "The Detection of Malignant Cells in Cytologic Material By Electronic Means - the Cytoanalyzer", in Proceedings of the Third National Cancer Conference, pp. 636-643, 1957.
- [206] Tolles, W.E., Horvath, W.J. and Bostrom, R.C., "A Study of the Quantitative Characteristics of Exfoliated Cells From the Female Genital Tract: I. Measurement Methods and Results", Cancer, Vol. 14, No. 3, 1961.
- [207] Tolles, W.E., Horvath, W.J., and Bostrom, R.C., "A Study of the Quantitative Characteristics of Exfoliated Cells From the Female Genital Tract: II. Suitability of Quantitative Cytological Measurements for Automatic Prescreening", Cancer, Vol. 14, No. 3, 1961.
- [208] Tolles, W.E., Mansberg, H.P., "Size and Shape Determination in Scanning Microscopy", Annals of the New York Academy of Science, Vol. 97, pp. 516-526, 1962.
- [209] Tweeddale, "The Reporting of Cytologic Smears: Selected Problems with Morphologic Terms", California Medicine, Vol. 109, pp. 13-14, 1968.
- [210] Underwood, E.E., Quantitative Stereology, Addison-Wesley, Reading, Massachusetts, 1970.
- [211] Van Dilla, M.A., et. al., "Cell Microfluorometry: A Method for Rapid Fluorescence Measurement", Science, Vol. 163, pp. 1213-1214, 1969.

- [212] Vastola, E.F., "Biological Image Processing in the Nervous System: Needs and Predictions", in Image Processing in Biological Science, Ransey, E. (Ed.), pp. 211-218, University of California Press, Berkeley, 1968.
- [213] Vincent Memorial Hospital Staff, The Cytologic Diagnosis of Cancer, W.B. Saunders Co., Philadelphia, Pennsylvania, 1950.
- [214] Von Haam, "An EDP Program for the Air Force Cytology Screening Project", Military Medicine, Vol. 135, pp. 574-578, 1970.
- [215] Voors, A.W., "Quantitative Evaluation of Cervical Cytology Screening Program", American Journal of Epidem., Vol. 90, pp. 214-223, 1969.
- [216] Wald, N. and Preston, K., "Automatic Screening of Metaphase Spreads for Chromosome Analysis", in Image Processing in Biological Science, Ransey, D. (Ed.), pp. 9-34, 1968.
- [217] Walton, W.H., "Survey of the Automatic Counting and Sizing of Particles", British Journal of Applied Physics, Supplement 3, pp. 121-125, 1954.
- [218] Watson, J.D., Molecular Biology of the Gene, Benjamin, New York, 1970.
- [219] Wheelless, L.L., et. al., "Computerized Cytofluorometric Research System", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 161-175, 1970.
- [220] Wied, G.L., et, al., "Fluorophotometric Analysis on Cervical Epithelial Cells", Acta Cytologica, pp. 61-67, 1964.
- [221] Wied, G.L., "Opinion Poll on Cytological Terminology", Acta Cytologica, Vol. 2, pp. 63-139, 1958.
- [222] Wied, G.L., "Opinion Poll on Cytological Definitions", Acta Cytologica, Vol. 2, pp. 26-62, 1958.
- [223] Wied, G.L., "Editorial: Computers or Cytotechnologists", Acta Cytologica, Vol. 12, pp. 1-2, 1968.
- [224] Wied, G.L., "Automated Cell Screening: A Practical Reality or a Daydream?", in Cytology Automation, Evans, D.M.D. (Ed.), pp. 43-47, Williams and Wilkins Co., Baltimore, 1970.
- [225] Wied, G.L. (ed.), "An International Agreement on Histological Terminology for Lesions of the Uterine Cervix", Acta Cytologica, Vol. 6, p. 235, 1962.
- [226] Wied, G.L., et. al., "Computer-Assisted Identification of Cells from Uterine Adenocarcinoma: A Clinical Feasibility Study with TICAS", Acta Cytologica, Vol. 12, pp. 357-370, 1968.
- [227] Wied, G.L., et. al., "Taxonomic Intra-Cellular Analytic System (TICAS) for Cell Identification", Acta Cytologica, Vol. 12, pp. 180-204, 1968.

- [228] Wied, G.L., et. al., "Computerized Recognition of Uterine Glandular Cells; II. The Application of a Self-Learning Program", *Acta Cytologica*, Vol. 13, pp. 662-671, 1969.
- [229] Wied, G.L., et. al., "Computer-Assisted Identification of Cells from Uterine Adenocarcinoma: A Clinical Feasibility Study with TICAS. II. Measurements at 590 NM.", *Acta Cytologica*, Vol. 13, pp. 21-26, 1969.
- [230] Wied, G.L., et. al., "Computer Discrimination of Blood Cells from Two Cases of 'Leukemic' Lymphosarcoma", *Acta Cytologica*, Vol. 13, pp. 688-695, 1969.
- [231] Wied, G.L., et. al., "Computerized Recognition of Uterine Glandular Cells: I. Assessment of a Self-Optimizing Recognition Program", *Acta Cytologica*, Vol. 13, pp. 611-619, 1969.
- [232] Wied, G.L., et. al., "TICAS Assessment of Cells From Atypical Hyperplasia of the Endometrium", *Acta Cytologica*, Vol. 13, pp. 552-556, 1969.
- [233] Wied, G.L., et. al., "Computerized Microdissection of Cellular Images", *Acta Cytologica*, Vol. 14, pp. 418-433, 1970.
- [234] Wied, G.L. and Bahr, G.F. (Eds.), Introduction to Quantitative Cytochemistry, Academic Press, New York, 1966.
- [235] Wied, G.L., Messina, A.M., and Rosenthal, E., "Comparative Quantitative DNA-Measurements on Feulgen Stained Cervical Epithelial Cells", *Acta Cytologica*, Vol. 10, pp. 31-37, 1966.
- [236] Wied, G.L., Bahr, G.F., and Bartels, P.H., "Laboratory Organization in the Detection of Early Cervical Neoplasia", *Obstetrical and Gynecological Survey*, Vol. 24, pp. 935-966, 1969.
- [237] Wied, G.L. and Bahr, G.F. (eds.), Introduction to Quantitative Cytochemistry - II., Academic Press, New York, 1970.
- [238] Wied, G.L., Bahr, G.F., and Bartels, P.H., "Automatic Analysis of Cell Images by TICAS", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 195-360, 1970.
- [239] Williams, H.P., "Application of Television Ultraviolet Microscope to the Direct Observation of Cytological Absorption Characteristics", *IRE Convention Record*, Vol. 3, p. 131.
- [240] Wolf, H.S., "An Apparatus for Counting Small Particles in Random Distribution with Special Reference to Red Blood Cell Corpuscles", *Nature*, Vol. 163, p. 967, 1950.
- [241] Yakovlev, P.I. and Sullivan, P.R., "Cell Count in the Lateral Geniculate Nucleus of Man", *Neurology*, Vol. 8, p. 566, 1958.

- [242] Young, I.T., "Automated Leukocyte Recognition", in Automated Cell Identification and Cell Sorting, Wied, G.L. and Bahr, G.F. (Eds.), pp. 187-194, 1970.
- [243] Young, J.Z. and Roberts, F., "High-Speed Counting With the Flying Spot Microscope", *Nature*, Vol. 169, p. 963, 1952.
- [244] Papanicolaou, G.N., "A New Procedure for Staining Vaginal Smears", *Science*, Vol. 95, pp. 438-439, 1942.
- [245] U.S. Naval Medical School, Manual for Exfoliative Cytology, National Naval Medical Center, Bethesda, Maryland.
- [246] Wied, B.L., et. al., "Cytology of Invasive Cervical Carcinoma and Carcinoma in Situ", *Annals of the New York Academy of Science*, Vol. 97, pp. 759-766, 1962.
- [247] Bohm, N., et. al., "A One-Parameter Approach to Cytology Automation by Rapid Reulgen Fluorescence Cytophotometry", in Cytology Automation, D.M.D. Evans (ed.), Livingston, Ltd., Edinburgh, 1970.
- [248] Gibbs, D.F., "Fluid Enzyme Methods in the Detection of Gynaecological Cancer", in Cytology Automation, D.M.D. Evans (ed.), Livingstone, Ltd., Edinburgh, 1970.
- [249] Ambrose, E.J., et. al., "Cell Agglutinins", in Cytology Automation, D.M.D. Evans (ed.), Livingstone, Ltd., Edinburgh, 1970.
- [250] Altmann, F.P., et. al., "Integrated Cellular Chemistry Applied to Malignant Cells", in Cytology Automation, D.M.D. Evans (ed.), Livingstone, Ltd., Edinburgh, 1970.
- [251] Ward, A. and McMaster, G.W., "Automatic Screening for Cervical Cancer", in Medical Computing.
- [252] Casperson, T. and Lomakka, "Scanning Microscopy Techniques", *Annals of the New York Academy of Science*, Vol. 97, pp. 449-463, 1962.
- [253] Tolles, W. (ed.), "Applications and Methods of Counting and Sizing in Medicine and Biology", *Annals of the New York Academy of Science*, Vol. 99, Article 2, June 1962.
- [254] Coons, A.H., Creech, H.J. and Jones, R.N., "Immunological Properties of an Antibody Containing a Fluorescent Group", *Proceedings of the Society of Experimental Biology and Medicine*, Vol. 47, pp. 200-202, 1941.
- [255] Fisher, C., "Analysing Images by Computer", *New Scientist and Science Journal*, September 23, 1971.
- [256] McCormick, B., "Advances in the Development of Image-Processing Hardware", in Image Processing in Biological Science, D. Ramsey (ed.), University of California Press, Los Angeles, 1968.

- [257] Unger, S.H., "A Computer Oriented Toward Spatial Problems", *Proceedings of the IRE*, Vol. 46, pp. 1744-1750, 1958.
- [258] Golay, M.J.E., "Hexagonal Parallel Pattern Transformations", *IEEE Transactions on Computers*, Vol. C-18, pp. 733-740, August 1962.
- [259] Huang, T.S., Schreiber, W.F., and Tretiak, O.J., "Image Processing", *Proceedings of the IEEE*, Vol. 59, no. 11, pp. 1586-1609, 1971.
- [260] Rutovitz, D., "Data Structures for Operations on Digital Images", in Pictorial Pattern Recognition, pp. 105-134, Thompson Book Co., Washington, 1968.
- [261] Husain, O.A.N. and Henderson, M., "Observations on the Use of the Quantimet Image Analysing Computer in Automatic Scanning for Malignant Cells", in Cytology Automation, D.M.D. Evans (ed.), pp. 214-227, Livingstone Ltd., Edinburgh, 1970.
- [262] Sweet, R.G., "High Frequency Recording With Electrostatically Deflected Ink Jets", *Review of Scientific Instruments*, Vol. 36, pp. 131-136, 1964.
- [264] Preston, K., Perkin-Elmer Optical Group release, CELLSCAN Status Report, May 1970.
- [265] Bartels, P. H., Bahr, G. F., and Wied, G. L. "Information Theoretic Approach to Cell Identification," *op.cit.* [242], pp. 361-384.
- [266] Houshovitz, S. B. "Autoradiographic Studies of Exfoliated Cells of the Uterine Cervix: A correlation with the Cytology Report," *Acta Cytologica*, Vol. 13, pp. 36-39, (1969).
- [267] Jagger, J., Introduction to Research in Ultraviolet Photobiology, Prentice-Hall, Englewood Cliffs, N.J., 1967.
- [268] Paul, J., Cell and Tissue Culture, fourth edition, Livingstone, Edinburgh, 1970.
- [270] Divilbiss, J. D. and McCormick, B. H., "Tentative Logical Realization of a Pattern Recognition Computer," Univ. of Ill. Digital Computer Laboratory File Note 403, October 10, 1961.
- [271] Spriggs, A. I., "Choice of Criteria for Screening," *loc.cit.* [46], pp. 24-27.
- [272] Dineen, G. P., "Programming Pattern Recognition," *Proceedings 1955 WJCC*, p. 94y, 1955.
- [273] Selfridge, O. G., "Pattern Recognition and Modern Computers," *Proceedings 1955 WJCC*, p. 91, 1955.
- [274] Rosenfield, A., and Pfaltz, T. C., "Sequential Operations in Digital Picture Processing," Journal of the ACM, Vol. 4, October 1966.

- [275] McCormick, B. H., et.al., "ILLIAC III Reference Manuals,"
Department of Computer Science Reports 433, 434, 472, 473,
University of Illinois, Urbana-Champaign, February-August 1971.
- [276] Castleman, K. R., et.al., "Method and Means for Chromosome
Karyotyping," Jet Propulsion Laboratory New Technology
Report No. 30-11852, Case No. 2500.
- [277] Butler, J. W., Butler, M. K., and Stroud, A., "Automatic
Classification of Chromosomes," Data Acquisition and
Processing in Biology and Medicine, Vol. 3, Pergamon Press,
New York, 1964.
- [278] Neurath, P. W., Bablouzian, B. L., et.al., "Human Chromosome
Analysis by Computer--an optical pattern recognition problem,"
Annals of the New York Academy of Sciences, Vol. 128,
pp. 1013-1028, 1966.
- [279] Center for Disease Control, U. S. Public Health Service,
"National Personnel Needs in Clinical Cytology Laboratories,"
Acta Cytologica, Vol. 16, No. 4, pp. 398-400, July-August 1972.

APPENDIX

BLOB

```

C      IPO NOW CONTAINS ONES FOR EACH PIXEL WITH
C      NO. OF HITS .GE. INTOL
CALL CLEAR(IPWK1,1)
CALL SHIFT(IPWK1,IPWK1,2*IRADUB,2*IRADUB)
CALL SHIFT(IPWK1,IPWK1,-IRADUB,-IRADUB)
CALL AND(IPC,IPO,IPWK1,0)
C      THE PREVIOUS FOUR INSTRUCTIONS ZEROED OUT THE
C      EFFECT OF THE BORDERS ON THE BLOB OPERATION.
DO 400 I=1,4
400   CALL AND(IPHITS(I),IPHITS(I),IPWK1,1)
      CALL RELFAS(14,IWKARA)
RETURN
300   WRITE(6,999)
399   FORMAT(1H1,'BLOB PARAMETER ERROR')
      CALL SNAP
END

```

AVG

```

SUBROUTINE AVG(/S4AVG/,/S4PICN/,KPCW)

```

```

C      SUBROUTINE TO AVERAGE OVR A SQUARE NEIGHBORHOOD IN
C      IN S4PICN OF SIZE 2** (KPCW*2). ANSWER IS TRUNCATED TO
C      FOUR BITS. CURRENTLY USES ROSENFELD'S ALGORITHM
C      (JEFF TRANS. ON COMPUTERS, C-20 NO. 5, PP. 562-569.)
C      KPCW MUST BE .LE. 4.

```

```

INTEGER S12SM1(12)/12*0/,S12SM2(12)/12*0/

```

```

C      I N I T I A L I Z E

```

```

C      IX=1
C      IY=1
C      KRITS=2*KPCW
C      CALL EQUALS(S12SM1,S4PICN,0,4)
C      CALL CLEARS(S12SM1(5),0,KRITS)

```

```

C      ++++++
C
C      J=KRITS+2
C      DO 10 KRITS=4,J,2
C      CALL SHIFTS(S12SM2,S12SM1,IX,0,KRITS)
C      CALL ADDSUB(S12SM1,KRITS+1,S12SM1,KRITS,S12SM2,KRITS,1)
C      CALL SHIFTS(S12SM2,S12SM1,0,IY,KRITS+1)
C      CALL ADDSUB(S12SM1,KRITS+2,S12SM2,KRITS+1,S12SM1,KRITS+1,1)
C      IX=IX+IX
C      IY=IY+IY
10   CONTINUE
C      KRITS=2*KPCW+4
C      CALL RELFAS(KRITS,S12SM2)
C      CALL EQUALS(S4AVG,S12SM1(KRITS-3),0,4)
C      CALL RELFAS(KRITS,S12SM1)
C      IX=-IX/4
C      IY=-IY/4
C      CALL SHIFTS(S4AVG,S4AVG,IX,IY,4)
C      RETURN
C      END

```

MAIN

```
INTEGER BP(4,3)/2,4,8,3,3,5,8,3,4,6,8,3/  
INTEGER PLNREF/0/,WIN(4)/1,1,128,128/,NRIG/3/  
INTEGER PLN(4)/4*0/,STKHIT(4)/4*0/,STKPIC(4)/4*0/  
CALL CLEAR(PLNREF,0)  
CALL FRAME(PLNREF,1)  
CALL CARDS(4,WIN,STKPIC)  
CALL CLEAR(PLN(4),0)  
CALL RESULT(STKPIC,0,4,0)  
DO 20 I=1,NRIG  
CALL BLGR(STKHIT,STKPIC,4,PLN(I),BP(1,I),BP(2,I),BP(3,I),BP(4,I))  
CALL CR(PLN(4),PLN(4),PLN(I),0)  
CALL PUNCH(4,WIN,PLN)  
CALL RESULT(PLN,PLNREF,4,8)  
RETURN  
END
```

LINES

```

SUBROUTINE LINES(/IDIFFS/,/IPO/,/IDIR/,/IDIST/,/ITHRES/)
  INTEGER FORTHQ(2,8)/0,-1,-1,-1,-1,0,-1,1,0,1,1,1,1,0,1,-1/,
1    RORTHQ(2,8) /0,1,1,1,1,0,1,-1,0,-1,-1,-1,-1,0,-1,1/,
2    ACCUM4(4)/4*0/,WK1/0/,WK2/0/,WK3/0/,
3    SHIFTR/0/
  CALL CLEAR4(ACCUM4,0,4)
  CALL EQUAL(SHIFTR,IDIFFS,0)
  DO 10 I=1,IDIST
    CALL SHIFT(SHIFTR,SHIFTR,FORTHQ(1,IDIR),FORTHQ(2,IDIR))
10   CALL ADD1TP(SHIFTR,ACCUM4,4)
    CALL SLICE(WK1,ACCUM4,4,4,ITHRES)
    CALL CLEAR(WK3,0)
    DO 20 I=1,IDIST
      CALL SHIFT(WK1,WK1,RORTHQ(1,IDIR),RORTHQ(2,IDIR))
20   CALL OR(WK3,WK2,WK1,0)
      CALL AND(IPQ,IDIFFS,WK3,0)
      CALL RELEAS(4,ACCUM4)
      CALL RELEAS(1,SHIFTR)
      CALL RELEAS(1,WK1)
      CALL RELEAS(1,WK2)
    RETURN
  END

```

BLK DATA

```

BLOCK DATA
COMMON /LS7/ LPR,LPA,LRR,LRE,LCE,LRA,FIRSTP,LASTP
COMMON /LS7S/ PSIZE,NUMSI7,LS(8)
DATA LASTP/45/ , NUMSI7/1/
INTEGER FIRSTP/1/,PSIZE/1/
DATA LS /1,2048,128,128,128,16,1,45/
COMMON /PLX/ NPLN,NBUF,NPU(2)
COMMON /WEP/ IW(4)
DATA LPR,LPA,LRR,LRE,LCE,LRA,NPLN,NBUF,IW
1/1,2048,128,128,128,16,45,45,1,1,128,128/
DATA NPU/2*0/
COMMON /PIX/ PLANES
COMMON /BUX/ JOLD,JMT,JNEW,BLOCK
COMMON /BUE/ BUFFFF
COMMON /RCW/ IR
  INTEGER PLANES(45) /45*0/, BLOCK(225),
1  IR(132),BUFFFF(23040) /3HPAX/
DATA JOLD,JMT,JNEW /3*0/
END

```

RIDGE

SUBROUTINE RIDGE(/S4DIFF/,/PCUT/,IDIR,PAVG)

SUBROUTINE TO FIND EDGE POINTS IN DIFFERENCED PICTURE.

1 INTEGER DIRS(2,8)/1,0,1,-1,0,-1,-1,-1,-1,0,-1,1,0,1,1,1/,
DIRSKL(17),PWK1/0/,PWK2/0/,S4WK1(4)/4*0/,PAVG

BUILD DIRECTION LIST

KAVG=2**PAVG

K=KAVG/2

DIRSKI(1)=KAVG

KX=DIRS(1,IDIR)

KY=DIRS(2,IDIR)

KKX=-K*KKX

KKY=-K*KKY

I=KAVG+1

I=2

DO 10 J=1,I

IF(J.EQ.K+1) GOTO 1

DIRSKI(I)=KKX

DIRSKI(J+1)=KKY

I=J+2

CONTINUE

KKX=KKX+KX

KKY=KKY+KY

CONTINUE

IDL=LOC(DIRSKI(1))

CALL MAXIMA(IDL,PWK1,S4DIFF,4)

CALL LINES(PWK1,PCUT,IDIR,4,2)

CALL RELEASE(1,PWK1)

RETURN

ENTRY DIRDIF(/S4OUT/,/S4IN/,PSIGN,IDIR,PAVG)

INTEGER PSIGN

KAVG=2**PAVG

KX=-KAVG*DIRS(1,IDIR)

KY=-KAVG*DIRS(2,IDIR)

CALL SHIFTS(S4WK1,S4IN,KX,KY,4)

CALL SUBSIN(S4OUT,4,PSIGN,S4IN,4,S4WK1,4)

CALL RELEASE(4,S4WK1)

RETURN

END

RACAR

SUBROUTINE PACAR(/SCIST/,/PGEDGE/,/PEDGE/,/PBLCB/,/IDIR/,/IDMAX/,

1 /ICRIG/)

PURPOSE: TO LOCATE ALL EDGES IN DIRECTION IDIR FROM POINTS
IN PBLCB AND TO RECORD THEIR DISTANCES IN SCIST. EDGES ARE
INDICATED BY CNFS IN PLANE PEDGE. SCIST HAS FOUR PLANES, SC
DISTANCES ARE QUANTIZED INTO RANGES OF IDMAX/4. EDGES
WHICH WERE FOUND ARE STORED IN PGEDGE (CCCC EDGES),

INTEGER SCIST(4),PGEDGE,PEDGE,PBLCB,PWK1/0/,PWK2/0/,PWK3/0/,

1 PWK4/0/,PWK5/0/,

2 DIRS(11,8)/5,1,C,2,0,2,1,2,-1,C,C,4,1,-1,2,-1,1,-2,C,C,0,0,

3 5,0,-1,0,-2,-1,-2,1,-2,0,0,4,-1,-1,-2,-1,-1,-2,0,C,0,0,0,

4 5,-1,0,-2,0,-2,1,-2,-1,C,0,4,-2,1,-1,1,-1,2,C,0,0,0,0,

5 5,0,1,-1,2,0,2,1,2,0,C,4,1,1,2,1,1,2,C,0,0,0/

CALL EQUAL(PWK1,PBLCB,0)

CALL CLEAR(PWK4,C)

CALL CLEAR(PWK5,C)

INCREM=IDMAX/4

K=IDIR+4

IF(K.GT.8) K=K-8

IDL2=LCC(DIRS(1,K))

IFL1=LCC(DIRS(1,IDIR))

DO 10 I=1,4

DO 20 J=1,INCREM

CALL MARK(IDL1,PWK1,PWK1,0)

CALL AND(PWK2,PWK1,PWK4,1)

CALL AND(PWK3,PWK2,PEDGE,0)

EDGES DETECTED ARE PWK2

CALL OR(PWK5,PWK5,PWK3,C)

CALL CCONNECT(IFL2,PWK2,PWK1,PWK3,JCCUNT)

CALL AND(SCIST(I),PWK3,PBLCB,0)

SCIST(I) CONTAINS PLCS AT DIST I*INCREM FROM EDGE

CALL EQUAL(PWK4,PWK1,0)

CONTINUE

CALL EQUAL(PGEDGE,PWK5,C)

CALL RELEAS(1,PWK1)

CALL RELEAS(1,PWK2)

CALL RELEAS(1,PWK3)

CALL RELEAS(1,PWK4)

CALL RELEAS(1,PWK5)

RETURN

END

ALCR

```
SUBROUTINE ALCR( IPHITS ,/IPI/,/NPI/,/IPO/,/IRADLR/,
1 /IRADUR/,/INTCL/,/IDRCP/)
```

```
-----
C      IPHITS  -STACK FOR ACCUMULATING HIT COUNTS
C      IPI     -INPUT STACK
C      NPI     -NO. OF PLANES IN INPUT STACK
C      IPO     -OUTPUT PLANE
C      IRADLR  -SEARCH RADIUS LOWER BOUND
C      IRADUR  -SEARCH RADIUS UPPER BOUND
C      INTCL   -MINIMUM NUMBER OF RADI WHICH MUST SUCCEED
C      IDRCP   -DROP IN GS TO CAUSE RADIUS SEARCH SUCCESS
C      -----
```

```
DIMENSION IPACC(4),IPSUR(4),ISHIF(16),IPHITS(4),IPDIFF(4)
```

```
DIMENSION IWNDOW(4),IWKARA(14)
```

```
DATA IWNDOW /1,1,128,128/,IWKARA/14*0/
```

```
EQUIVALENCE (IPACC(1),IWKARA(1)),(IPSUR(1),IWKARA(5)),
```

```
C (IPWK1,IWKARA(9)),(IPDIFF(1),IWKARA(10)),(IPTCCC,IWKARA(14))
```

```
REAL FACTOR (8)/.59,.73,.59,1.0,.59,.73,.59,1.0/
```

```
INTEGER IDISTI(8),IDISTU(8)
```

```
DATA ISHIF /-1,-1,0,-1,1,-1,1,0,1,1,0,1,-1,1,-1,0/
```

```
C      BUILD A TABLE OF SEARCH DISTANCES, CORRECTED FOR A
C      RECTANGULAR RASTER HAVING AN ASPECT RATIO OF 1.42:1.95.
```

```
DO 100 I=1,8
```

```
  IDISTI(I)=FLOAT(IRADLR)*FACTOR (I)+.5
```

```
100  IDISTU(I)=FLOAT(IRADUR)*FACTOR (I)+.5
```

```
  IF(IRADLR.LT.1)GOTO 900
```

```
  IF(IRADLR.GT.IRADUR) GOTO 900
```

```
CALL CLEARS(IPHITS,0,4)
```

```
  IDIR=1
```

```
DO 300 I=1,15,2
```

```
  IRIR=IDISTI(IDIR)
```

```
  IRUR=IDISTU(IDIR)
```

```
  CALL EQUALS(IPSUR,IPI,0,4)
```

```
  CALL EQUALS(IPACC,IPI,0,4)
```

```
  IX=ISHIF(I)
```

```
  IY=ISHIF(I+1)
```

```
DO 200 ITERAT=1,IRUR
```

```
  IF(ITERAT.NE.IRIR)GOTO 20
```

```
  CALL SUBSIN(IPDIFF,4,IPWK1,IPI,4,IPACC,4)
```

```
  CALL SLICE(IPTCCC,IPDIFF,4,4,IDRCP)
```

```
20  CALL SHIFTS(IPSUR,IPSUR,IX,IY,4)
```

```
  CALL COMPER(IPWK1,IPSUR,IPACC,4,1)
```

```
  CALL COMBIN(IPACC,IPACC,IPSUR,4,IPWK1)
```

```
300  CONTINUE
```

```
  CALL SUBSIN(IPDIFF,4,IPWK1,IPI,4,IPACC,4)
```

```
  CALL SLICE(IPWK1,IPDIFF,4,4,IDRCP)
```

```
  CALL AND(IPWK1,IPWK1,IPTCCC,1)
```

```
  CALL ADDITP(IPWK1,IPHITS,4)
```

```
  IDIR=IDIR+1
```

```
300  CONTINUE
```

```
CALL SLICE(IPO,IPHITS,4,4,INTCL)
```

U. S. ATOMIC ENERGY COMMISSION
UNIVERSITY-TYPE CONTRACTOR'S RECOMMENDATION FOR
DISPOSITION OF SCIENTIFIC AND TECHNICAL DOCUMENT

(See Instructions on Reverse Side)

1. AEC REPORT NO. COO-2118-0029 UIUCDCS-R-72-497	2. TITLE Parallel Image-Processing for Automated Cytology
--	--

3. TYPE OF DOCUMENT (Check one):

- ☒ a. Scientific and technical report
- ☐ b. Conference paper not to be published in a journal:
- Title of conference _____
- Date of conference _____
- Exact location of conference _____
- Sponsoring organization _____
- ☐ c. Other (Specify) _____

4. RECOMMENDED ANNOUNCEMENT AND DISTRIBUTION (Check one):

- ☒ a. AEC's normal announcement and distribution procedures may be followed.
- ☐ b. Make available only within AEC and to AEC contractors and other U.S. Government agencies and their contractors.
- ☐ c. Make no announcement or distribution.

5. REASON FOR RECOMMENDED RESTRICTIONS:

6. SUBMITTED BY: NAME AND POSITION (Please print or type)

John S. Read
Research Programmer

Organization
Dept. of Computer Science
University of Illinois
Urbana, Illinois 61801

Signature 	Date October 1972
--	----------------------

FOR AEC USE ONLY

7. AEC CONTRACT ADMINISTRATOR'S COMMENTS, IF ANY, ON ABOVE ANNOUNCEMENT AND DISTRIBUTION RECOMMENDATION:

8. PATENT CLEARANCE:

- ☐ a. AEC patent clearance has been granted by responsible AEC patent group.
- ☐ b. Report has been sent to responsible AEC patent group for clearance.
- ☐ c. Patent clearance not required.

BIBLIOGRAPHIC DATA SHEET		1. Report No. UIUCDCS-R-72-497	2.	3. Recipient's Accession No.
1. Title and Subtitle Image Processing for Automated Cytology			5. Report Date June, 1972	
			6.	
7. Author(s) John Stevenson Read			8. Performing Organization Rept. No.	
9. Performing Organization Name and Address Dept. of Computer Science Univ. of Ill. at Urbana-Champaign Urbana, Illinois 61801			10. Project/Task/Work Unit No. ILLIAC III	
			11. Contract/Grant No. AT(11-1)-2118	
12. Sponsoring Organization Name and Address U.S. Atomic Energy Commission			13. Type of Report & Period Covered	
			14. Feb.-June 1972	
15. Supplementary Notes				
16. Abstracts <p>The prime objective of this work was to examine the effectiveness of a parallel digital image processor in the analysis of cervical (Pap) smear imagery. This task has historically been shown to be very difficult to do by machine, partly because clumping and overlapping of cells has presented difficulties exceeding the capacity of cost-effective image-processing technology. The ILLIAC III's special parallel image processor promises a capability to analyze much more complex patterns for a given cost than has been possible in the past.</p> <p>This thesis contains background information on cervical smears, an extensive review of automated cytology, and reports on three experiments: in (1) detecting enlarged cell nuclei in the presence of clumps of leukocytes, (2) detecting superficial epithelial cells, and (3) discriminating two important textures occurring in cervical preparations. The last experiment utilized the Varivalued logic approach to pattern recognition developed by McCormick and Michalski.</p>				
17. Key Words and Document Analysis. 17a. Descriptors Digital Image processing Automated cytology Parallel processors Biological cells Automated Cancer screening ILLIAC III Pattern Articulation Unit Varivalued logic Texture analysis Image segmentation 17b. Identifiers/Open-Ended Terms				
17c. COSATI Field/Group				
18. Availability Statement Release Unlimited			19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages
			20. Security Class (This Page) UNCLASSIFIED	22. Price

NOV 29 1972

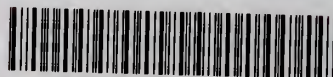
JUL 26 1973



UNIVERSITY OF ILLINOIS-URBANA

510.84 IL6R no. C002 no. 493-498(1972)

Tree height reduction for parallel proce



3 0112 088400210